

STIC Biotech/ChemLib

76549

Fr m: Rao, Manjunath N.
Sent: Thursday, September 26, 2002 11:05 AM
To: STIC-Biotech/ChemLib
Subject: Sequence search request for 09/870,113

From: Manjunath N. Rao
Art Unit 1652, Room 10A11
Mail Box in Room 10C 01
Phone: 306-5681

CRIFE

Date: 9-20-02

Please search the following as soon as possible for application with serial number **09/870,113**

SEQ ID NO:3, 5,7,9,11, against all commercial nucleic acid databases including issued patents database and pending application database and provide a **print of all results**.

If y u have any questions please call me at the above phone number.

Thanks

Edward Hart
Technical Info. Specialist
STIC/Biotech
CMI 6B02 Tel: 305-9203

Manjunath N. Rao, Ph.D.
Biotechnology Patent Examiner
Art Unit 1652, Room 10A11
Crystal Mall 1, USPTO.

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 9/26/02
Date Completed: 9/29/02
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH: 5
NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: PS
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2002, 03:51:19 ; Search time 5287.54 Seconds
(without alignments)
2303.386 Million cell updates/sec

Title: US-09-870-113-3
Perfect score: 582
Sequence: 1 atggagttgagggcgggg.....tgaaacctgcgaagctga 582

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg_hum.*
- 31: em.htg_inv.*
- 32: em.htg_other.*
- 33: em.htgo_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
------------	-------	-------	-------	--------	----	-------------

1	577	99.1	1448	9	AF327402	AF327402	Homo sapi	
2	577	99.1	1448	17	HSA303077	AJ303077	Homo sapi	
3	520.4	89.4	1889	9	AF327403	AF327403	Homo sapi	
4	520.4	89.4	1889	17	HSA303078	AJ303078	Homo sapi	
5	454.4	78.1	1322	6	AX061229	AX061229	Sequence	
c	291	50.0	123160	9	AL353719	AL353719	Human DNA	
7	273	46.9	401	6	AX071440	AX071440	Sequence	
8	251	43.1	221062	2	AC096351	AC096351	Rattus no	
c	9	210.6	36.2	221062	2	AC096351	Rattus no	
10	207.2	35.6	4112	10	AF288621	AF288621	Mus muscu	
11	198.8	34.2	1429	9	AF223466	AF223466	Homo sapi	
12	177.8	30.5	5869	10	AF361699	AF361699	Mus muscu	
13	174.2	29.9	3415	9	AY032628	AY032628	Homo sapi	
14	172	29.6	858	9	BC015013	BC015013	Homo sapi	
15	125.2	21.5	163205	2	AC099368	AC099368	Rattus no	
16	116.2	20.0	191702	2	AC108878	AC108878	Mus muscu	
17	113.8	19.6	108765	2	AF216674	AF216674	Homo sapi	
c	18	112.2	19.3	174445	9	AC051642	Homo sapi	
c	19	112.2	19.3	182504	2	AC022597	Homo sapi	
20	105.2	18.1	483	6	AX306258	AX306258	Sequence	
c	21	82.4	14.2	34305	3	CEM02B12	266521	Caenorhabdi
22	72.4	12.4	3037	3	AY060268	AY060268	Drosophil	
c	23	71.6	12.3	146743	2	AF003518	Oryza sat	
24	70.8	12.2	2878	3	AF217402	AF217402	Drosophil	
c	25	68	11.7	615	11	CNS06FGL	AL396379	T7 end of
26	68	11.7	71053	2	AC102327	AC102327	Mus muscu	
c	27	67.2	11.5	77191	2	AC023381	AC023327	Homo sapi
28	67.2	11.5	207420	2	AC078884	AC078884	Mus muscu	
29	67	11.5	121146	2	AC108558	AC108558	Rattus no	
30	66.6	11.4	38723	2	AC095142	AC095142	Rattus no	
31	66.6	11.4	59632	2	AC102213	AC102213	Mus muscu	
32	66.6	11.4	94914	2	AC105592	AC105592	Rattus no	
33	66.2	11.4	158167	2	AC034105	AC034105	Homo sapi	
34	66	11.3	65928	2	AC012570	AC012570	Homo sapi	
35	65.8	11.3	79355	2	AC021280	AC021280	Homo sapi	
c	36	65.6	11.3	46181	2	AC108555	AC105855	Rattus no
37	65.6	11.3	115715	2	AC079411	AC079411	Homo sapi	
38	65.6	11.3	140469	2	AC025066	AC025066	Homo sapi	
39	65.6	11.3	148407	2	AC106354	AC106354	Rattus no	
-	40	65.2	11.2	1560	14	HS21E	M29384	Herpes simp
c	41	65.2	11.2	62649	2	AC022552	AC022552	Homo sapi
42	65.2	11.2	146744	2	AC098461	AC098461	Rattus no	
43	65.2	11.2	174096	2	AC097177	AC097177	Rattus no	
44	65	11.2	93663	2	AC107039	AC107039	Pan trogl	
45	64.8	11.1	60984	2	AC102552	AC102552	Mus muscu	

ALIGNMENTS

RESULT	1
AF327402	
LOCUS	AF327402 1448 bp mRNA linear PRI 02-MAY-2001
DEFINITION	Homo sapiens putative mitochondrial solute carrier splice variant mRNA, complete cds, alternatively spliced, nuclear gene for mitochondrial product.
ACCESSION	AF327402
VERSION	AF327402.1 GI:13926046
KEYWORDS	human.
SOURCE	Homo sapiens
ORGANISM	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 1448)
AUTHORS	Li.F.Y., Nikali.K., Grogan.J., Leibiger.I., Leibiger.B., Schweyen.R., Larsson.C. and Suomalainen.A.
TITLE	Characterization of a novel human putative mitochondrial transporter homologous to the yeast mitochondrial RNA splicing proteins 3 and 4
JOURNAL	FEBS Lett. 494 (1-2), 79-84 (2001)
MEDLINE	21195335
PUBMED	11297739
REFERENCE	2 (bases 1 to 1448)

```

AUTHORS      Li,F., Nikali,K., Gregan,J., Leibiger,I., Leibiger,B., Schweyen,R.,
              Larsson,C. and Suomalainen,A.
TITLE        Direct Submission
JOURNAL      Submitted (08-DEC-2000) Human Molecular Genetics, National Public
              Health Institute, Mannerheimintie 166, Helsinki 00300, Finland
FEATURES
  source      1. .1448
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /chromosome="10"
              /map="10q24"
              30. .1124
              /note="similar to mitochondrial RNA splicing protein 3/4;
              alternatively spliced"
              /codon_start=1
              /product="putative mitochondrial solute carrier splice
              variant"
              /protein_id="AAK49519.1"
              /db_xref="GI:13926047"
              /translation="MELEGRGAGGVAGGPAAGPGRSGESALLDGLQRCVGRGAGGG
              EGACAPPVRQDPSGDPEALPAGATVTHMVGAVAGILEHCVMYPIDCVKTRMOS
              LDTPDAARYRNLEALWRIITRGLWRPRLGNVTATGAGPAHALYFACYEKLKRLIS
              LQPHGNSHIANAGACVATLLHDAAMNPAEVKQRMOMYNSPYHRVTD CVRAVMQN
              EGAGAFVRSVTTOLTWNVPQATHEMTYFLOEHENPQREYNPSHHVLSGACAGAVAA
              AATPLDVCKTLNTQESALNHSHTGHTCMASAFRTVYQVGCVTAYFRFGQARVY
              QIPSTAIAVSVEFFELITKTRQEEWRAGK"
              322 a 408 c 435 g 283 t
BASE COUNT
ORIGIN

```

Query Match	99.1%	Score 577:	DB 9:	Length 1448:
Best Local Similarity	100.0%:	Pred. No. 3.le-83:		
Matches 577:	Conservative	0:	Mismatches	0:
			Indels	Gaps
QY	1	atgagtgtagggcggggctgctggcggtgtgagcgggggcgagcgagggcccggg	60	
DB	30	ATGAGTTTGAGGGCGGGGTGCTGGCGGTGTGCGGGGGGGCGCGGAGGCCCGGG	89	
QY	61	cgagccccgggagtcggcgctgtgagcgggtggctgcagcgggcgctgggcccgggg	120	
DB	90	CGGAGCCCGGGAGTCGGCGCTGCTGGACGGGTGGCTGCAGCGGGCGTGGGCGGGG	149	
QY	121	gcgcgcggggagggcggggctgcaggcccccggtacgacaaagatcgggactcgggc	180	
DB	150	GC CGCGGGCGGGAGGCGGGGGCTTGCAGGCCCGCGTACGACAAATCCGGATCCGGC	209	
QY	181	ccggactcacgggcgctgcgcgctggagccactgtcaccacgcacatggtggcagcgcc	240	
DB	210	CCGGACTACGAGCGCTGCCGGCTGGAGCCACTGTCAACACGCACATGTTGGCAGGGGCC	269	
QY	241	gtggcagggtactgtgagcactcgtgtgatatcccatcactcgtcgaagaccggatg	300	
DB	270	GTGGCAGGGATCTGGAGCACTCGTGTATGCCCATGCCATCGCTCAAGACCGGATG	329	
QY	301	cagagttacagctcgtaccagctgcccgctatcgcaatgtgtggagggccctctggag	360	
DB	330	CAGAGCTACAGCTGACCCAGCTGCCCGTATCGCAATGTGTTGGAGGGCCCTCTGAGG	389	
QY	361	attataagaacgagggccctatggagggcccatgaggggctgaacgtcacagacaagcc	420	
DB	390	ATTATAAGAACGGAGGGCCCTATGGAGGCCCATGAGGGGGCTGAACGTACAGCAACAGGC	449	
QY	421	gcagggcctgccacgccccttatttggctgtctagaaaaagttaaaaaagacattgagt	480	
DB	450	GCAGGGCCTCCCAAGCCCTTTATTTGGCTGTCTACGAAAGTTAAAAAGACATTGAGT	509	
QY	481	gatgaatccaccctgggggcaatagccatattgccaatgttcggcccgaggtgtggcca	540	
DB	510	GATGTAAATCCACCCTGGGGGCATAGGCATATTGGCCNATGTCGGCGCGGTGTGGCCA	569	
QY	541	acattacttcattgatcgagccatgaaccctgcggaag	577	
DB	570	ACATTTACTTCATGATCAGCGCATGAACCCCTGCGGAAG	606	

RESULT 2

```

HSA303077 standard; RNA; HUM; 1448 BP.
XX HSA303077
XX AJ303077;
XX SV AJ303077.1
XX 02-FEB-2001 (Rel. 66, Created)
XX 26-APR-2001 (Rel. 67, Last updated, Version 2)
XX Homo sapiens mRNA for mitochondrial RNA splicing protein 3/4 (HMRS3/4
DE gene), 1448 bp cDNA splice variant
XX HMRS3/4 gene; mitochondrial RNA splicing protein 3/4.
XX Homo sapiens (human)
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
OC Eutheria; Primates; Catarrhini; Homnidae; Homo.
RN [1]
RP 1-1448
RA Nikali K.;
RT ;
RL Submitted (12-DEC-2000) to the EMBL/GenBank/DBJ databases.
RL Nikali K., Human Molecular Genetics, National Public Health Institute,
RL Mannerheimintie 166, Helsinki, FI-00300, FINLAND.
XX [2]
RA Li F., Nikali K., Gregan J., Leibiger I., Leibiger B., Schweyen R.,
RA Larsson C., Suomalainen A.;
RA "Characterization of a novel human putative mitochondrial transporter
RT homologous to the yeast mitochondrial RNA splicing proteins 3 and 4";
RL FEBS Lett. 494:79-84(2001).
XX key Location/Qualifiers
FH source 1..1448
FT /db_xref="taxon:9606"
FT /organism="Homo sapiens"
CDS 30..1124
FT /gene="HMRS3/4"
FT /product="mitochondrial RNA splicing protein 3/4"
FT /function="putative mitochondrial solute carrier"
FT /protein_id="CAC27936.1"
FT /translation="MELEGRGAGVAGGPAAGPGSPCESALLDGLQGVGRGAGGGE
FT AGACRPVPDPDPSGPDYIIRTPAGATVTTHMAGAVAGILEHCVMYPIDCVKTRMQSLQ
FT PGPAARYRVNLEALWRIIRT EGLMRPMLGNLTATGAGPAHALYFACYEKUKKTLSDI
FT HPGNSHIANGAAGCATLLHDAMNPAAEVVKROMYNTSPYHRVTCBRAVWONEGAG
FT AFYSXYTLTMVPQOAHFMPTVEFLQEHENPORRYNPSSHVLSGCAGAAAAATTP
FT LDVCKTLLNTQMSALNHSHTGITGTMASAFRTVYQGGVTAYFRGVAQRIYQIPSTA
FT IAWSVYEFFYLITKKREWRAGK"
XX Sequence 1448 BP: 322 A; 408 C; 435 G; 283 T; 0 other;
```

Query Match	99.1%	Score 577;	DB 17;	Length 1448;
Best Local Similarity	100.0%;	Pred. No. 3.1e-83;		
Matches 577; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	atggagttggaaggccgggtgctggcggtagtgcggggggggcgccggcggcaggccccggg	60
Dd	30	ATGGAGTTGAGGGCGGGGTCTGGCGGCTGTGC GGCGGGGCCGGCGCAGGGCCCCGGG	89
QY	61	cgaagcccggggagtcggcgctactgagacgggtggtctgcagcggggcgtgagggcggggg	120
Dd	90	CGAGACCCCGGGGAGTTCGGCGCTCTGTCAGCGGGTGCTTCAGCGGGGGCGTGGGCCGGGGG	149
QY	121	gccggcgccgggaggccggggcctgcaggccccccggtcacacaagaatcccgactccggc	180
Dd	150	GCGGGCGGGGAGCGCGGGGCTGTCAAGCCCCCGGTAACGAAGAATATCCGACATCCGCGC	209


```

Rattus.
1 (bases 1 to 221062)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbaria,J.,
Benton,J., Bimege,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowie,S., Brieva,M., Brown,M., Bryant,N.P., Buhay,C.,
Burck,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
Hollins,B., Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M.,
Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,
Ogoh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,
Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N.,
Sisson,I., Sodergren,E., Sonaik,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
Watlington,S., Williams,G., Williamson,A., Wleczkyk,R., Wooden,S.,
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 221062)
Worley,K.C.
Direct Submission
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Dec 20, 2001 this sequence version replaced gi:15627972.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GEZD
Center clone name: CH230-24M6
----- Summary Statistics
Assembly program: Phrap; version 0.990329First call to
findPhrapList
Consensus quality: 200164 bases at least Q30
Consensus quality: 205303 bases at least Q40
Consensus quality: 209366 bases at least Q20
Estimated insert size: 209880; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-gel estimation
Quality coverage: 3.9x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 33 contigs. The true order of the pieces
* is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* As soon as it is available and the accession number will
* be preserved.
1
32851: contig of 32851 bp in length
32951: gap of unknown length
32952: contig of 16033 bp in length
48984: gap of unknown length
48985: gap of unknown length
49085: contig of 14731 bp in length
63815: gap of unknown length
63916: contig of 12090 bp in length
76005: gap of unknown length
76105: contig of 12736 bp in length
88841: gap of unknown length
88842: contig of 11650 bp in length
100591: gap of unknown length
100592: contig of 11011 bp in length
111702: gap of unknown length
111703: contig of 10862 bp in length
122665: gap of unknown length
122765: contig of 6947 bp in length
129712: gap of unknown length
129811: contig of 7436 bp in length
137247: gap of unknown length
137248: contig of 7329 bp in length
144676: gap of unknown length
144777: contig of 7348 bp in length
152124: gap of unknown length
152225: contig of 6654 bp in length
158878: gap of unknown length
158879: contig of 4950 bp in length
163928: gap of unknown length
163929: contig of 5325 bp in length
164028: gap of unknown length
164029: contig of 6137 bp in length
169353: gap of unknown length
169354: contig of 4785 bp in length
175590: gap of unknown length
175591: contig of 4082 bp in length
180476: gap of unknown length
180477: contig of 5575 bp in length
184558: gap of unknown length
184559: contig of 3702 bp in length
190333: gap of unknown length
190334: contig of 2358 bp in length
194135: gap of unknown length
194136: contig of 3838 bp in length
196593: gap of unknown length
196594: contig of 2179 bp in length
200531: gap of unknown length
200532: contig of 2125 bp in length
202810: gap of unknown length
202811: contig of 2042 bp in length
205035: gap of unknown length
205036: contig of 1569 bp in length
207177: gap of unknown length
207178: contig of 2326 bp in length
208846: gap of unknown length
208847: contig of 1385 bp in length
211272: gap of unknown length
211273: contig of 1931 bp in length
213302: gap of unknown length
213303: contig of 2082 bp in length
215485: gap of unknown length
215486: contig of 1385 bp in length
216959: gap of unknown length
216960: contig of 1289 bp in length
218358: gap of unknown length
218359: contig of 1428 bp in length
219887: gap of unknown length
219888: contig of 1076 bp in length.
Location/Qualifiers
1. 221062
/organism="Rattus norvegicus"
FEATURES
source

```

```

/db_xref="taxon:10116"
/chromosome="Rf1"
/clones="CH230-24M6"
BASE COUNT      58389 a 50577 c 49732 g 58877 t 3487 others
ORIGIN

Query Match      43.1%; Score 251; DB 2; Length 221062;
Best Local Similarity 91.4%; Pred. NO. 1.9e-31;
Matches 266; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Qy 1 atgaagttggagggcggtgctgcgtgctggtggtggtggggggcgccgagggcccgccg 60
|||||
Db 72865 ATGAGATTGGAGGGCGGAGTGCAGCGCGGTGGCGGGAGGACCGACTGCTGGCGCCGGG 72924
|||||

Qy 61 cggagccccgggagtcgctgctgctgacggtggtgctgcagcggtggtggtggccggggg 120
|||||
Db 72925 CGGAGCCCCGGGAGTGCCTGCTGTGACGCGGTGGCTGCAGCGGGGCGTGGCGCGGGG 72984
|||||

Qy 121 gccggcgggggggggcgccgggctgcagggccggccgggtacacgaagatccggaactccggc 180
|||||
Db 72985 GCCGGCGGGGGAGCGCGGGGGCGCTATCAGACACCGCTGTACGCTGGATCCGAGTCCGGT 73044
|||||

Qy 181 ccgactacagggcgctccggctggagccactgacacacacacacacacacacacacacacac 240
|||||
Db 73045 CCGGAATACGAAGCGCTCCCGGCTGGAGCCACTGTGCACCGCACATGGTGGCGGGCGCC 73104
|||||

Qy 241 gtggcagggtacctggagcactgcgtggtggtgaccccatcgactgcgtcaag 291
|||||
Db 73105 GTGGCAGGATCCTGGAGCATGCGTGTGTACCCGATCGACTGCGTCAAG 73155
|||||

RESULT 9
AC096351/c
LOCUS
DEFINITION Rattus norvegicus chromosome Rf1 clone CH230-24M6, WORKING DRAFT
SEQUENCE, 33 unordered pieces.
AC096351
VERSION AC096351.2 GI:17944054
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 221062)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., AmaralTunge,H.C., Are,J.R., Banks,T.T., Barbara,J.,
Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
Hollins,B., Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapa,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M.,
Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokkwo,S.,

```

```

Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,
Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshitari,N.,
Sisson,I., Sodergren,E., Sonaik,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
Watlington,S., Williams,G., Williamson,A., Wleczek,R., Wooden,S.,
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 221062)
Worley,K.C.
Direct Submission
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Dec 20, 2001 this sequence version replaced gi:15627972.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GE2D
Center clone name: CH230-24M6
----- Summary Statistics
Assembly program: Phrap; version 0.990329First call to
findPhrapList
Consensus quality: 200164 bases at least Q40
Consensus quality: 205303 bases at least Q30
Consensus quality: 209366 bases at least Q20
Estimated insert size: 209880; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 3.9x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 33 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1
* 32851: contig of 32851 bp in length
* 32852: gap of unknown length
* 32952: contig of 16033 bp in length
* 48984: gap of unknown length
* 48985: contig of 14731 bp in length
* 63815: gap of unknown length
* 63915: contig of 12090 bp in length
* 76005: gap of unknown length
* 76106: contig of 12736 bp in length
* 88841: gap of unknown length
* 88941: contig of 11650 bp in length
* 100591: gap of unknown length
* 100592: contig of 11011 bp in length
* 111702: gap of unknown length
* 111802: contig of 10862 bp in length
* 122664: gap of unknown length
* 122764: contig of 6947 bp in length
* 122765: gap of unknown length
* 129711: contig of 7436 bp in length
* 129811: gap of unknown length
* 137247: contig of 7329 bp in length
* 137248: gap of unknown length
* 137348: contig of 7348 bp in length
* 144677: gap of unknown length
* 144777: contig of 7348 bp in length
* 152125: gap of unknown length

```

```

* 152225 158878: contig of 6654 bp in length
* 158879 158978: gap of unknown length
* 158979 163928: contig of 4950 bp in length
* 163929 164028: gap of unknown length
* 164029 169353: contig of 5325 bp in length
* 169354 169453: gap of unknown length
* 169454 175590: contig of 6137 bp in length
* 175591 175690: gap of unknown length
* 175691 180475: contig of 4785 bp in length
* 180476 180575: gap of unknown length
* 180576 184657: contig of 4082 bp in length
* 184658 184757: gap of unknown length
* 184758 190332: contig of 5575 bp in length
* 190333 190432: gap of unknown length
* 190433 194134: contig of 3702 bp in length
* 194135 194234: gap of unknown length
* 194235 196592: contig of 2358 bp in length
* 196593 196692: gap of unknown length
* 200530 200530: contig of 3838 bp in length
* 200531 200630: gap of unknown length
* 200631 202809: contig of 2179 bp in length
* 202810 202909: gap of unknown length
* 202910 205034: contig of 2125 bp in length
* 205035 205134: gap of unknown length
* 205135 207176: contig of 2042 bp in length
* 207177 207276: gap of unknown length
* 207277 208845: contig of 1569 bp in length
* 208846 208945: gap of unknown length
* 208946 211271: contig of 2326 bp in length
* 211272 211371: gap of unknown length
* 211372 213302: contig of 1931 bp in length
* 213303 213402: gap of unknown length
* 213403 215484: contig of 2082 bp in length
* 215485 215584: gap of unknown length
* 215585 216969: contig of 1385 bp in length
* 216970 217069: gap of unknown length
* 217070 218358: contig of 1289 bp in length
* 218359 218458: gap of unknown length
* 218459 219886: contig of 1428 bp in length
* 219887 219986: gap of unknown length
* 219987 221062: contig of 1076 bp in length.
FEATURES
  source
    Location/Qualifiers
      1. .221062
        /organism="Rattus norvegicus"
        /db_xref="taxon:10116"
        /chromosome="Rf1"
        /clone="CH230-24W6"
BASE COUNT  58389 a 50577 c 49732 g 58877 t 3487 others
ORIGIN

```

```

Query Match      36.2%; Score 210.6; DB 2; Length 221062;
Best Local Similarity 94.0%; Pred. No. 5.5e-25;
Matches 219; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 290 agaccggatgcagagctacagcctgaccagctgcccgtatcgcaatgtgttgaggg 349
Db 50109 AGACCCGGATGCAGAGCTACAGCCTATCCAGCGCCGCCGTATCGAAGCGTGTGGAGG 50050

QY 350 ccctctgagagattataagaacaggaggccctatgaggcccatgaggggctgaacgtca 409
Db 50049 CTCCTCGAGAAATTATAGAACAAGAGGGCCCTGTGTGAGGGCCCATGAGGGGCTGAACGTCA 49990

QY 410 cagcaacaggcggagcctgcccacgcccctttatttctgctacgaaaagttaaaaa 469
Db 49989 CTGGCACTGGCGGGGGCCCTGCCACGCCCTCTATTATTTTGGCTGCTACAAAAGTTAAAAA 49930

QY 470 agacattgagtgtatgaatcacacctgggggcaatagccattatgccaatggt 522
Db 49929 AGACATTGAGTGATGTAATCCACCCCGGGGCAATAGCCATATTGCGCAATGGT 49877

RESULT 10

```

```

AF288621
LOCUS      4112 bp      mRNA      linear      ROD 27-OCT-2001
DEFINITION Mus musculus mitochondrial solute carrier-like protein mRNA, complete cds.
ACCESSION AF288621
VERSION   GI:16506177
KEYWORDS  SOURCE
ORGANISM  house mouse.
           Mus musculus
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 4112)
AUTHORS   Li,Q.-Z., Ruan,Q.-G., Eckenrode,S., Shi,J.-D., Cruz,P., Wang,C.-Y. and She,J.-X.
TITLE      A new gene which is highly expressed in NOD mice spleen
JOURNAL    Unpublished
AUTHORS    2 (bases 1 to 4112)
           Li,Q.-Z., Shi,J.-D., Cruz,P., Echenrode,S., Ruan,Q.-G. and Wang,C.-Y.
DIRECT SUBMISSION
JOURNAL    Submitted (20-JUL-2000) Pathology, University of Florida, 1600 SW Archer Road, Rm. D6-15, Gainesville, FL 32610, USA
FEATURES
  source
    Location/Qualifiers
      1. .4112
        /organism="Mus musculus"
        /db_xref="taxon:10090"
        116..1132
        /note="50E12"
        /codon_start=1
        /product="mitochondrial solute carrier-like protein"
        /protein_id="AAL23859.1"
        /db_xref="GI:16506178"
        /translation="MELRRGGVGNQAARRMDGDRDGGGSKDAGSEYENLPTSASVSTMTAGAMAGILEHSIMVPDVSVKRMQSLNDPKARYTSIYGALKRIMTTEGFWRPLRGLNVMGMGAGPAHAMYFACVYENMKRTLNDVFSHQSHLANGVSGWATLLHDAVMNPAEVVKORLQMYNSQHSASRSCIRTVWRTGLGAFYRSYITQLTNNIPFQSIHFITYEFLQGVNPRDIPQSHIISGLAGALAAATPDVCKTLTLLNTOENMALNANVSGRISGMANAFRTVIYQLNLGAFYKQARVIYQMPSTAIYSWSVYEFKYLTKRQLENRTLY"
BASE COUNT  970 a 1024 c 970 g 1148 t
ORIGIN

```

```

Query Match      35.6%; Score 207.2; DB 10; Length 4112;
Best Local Similarity 64.4%; Pred. No. 3.7e-24;
Matches 327; Conservative 0; Mismatches 178; Indels 3; Gaps 1;

QY 73 gagtccggcgtctgacgggtggtgctgcagcggggcgctggccggggcgccgcgagg 132
Db 104 CGGGGGCGCGGATGGAGCTCCGCCGTGGCGGCTGGGGGAACAGGCTGGCGCGCGGAGG 163

QY 133 gaggccggggcctgcagggcccccggtacgacaagatccggaactccggccc---ggactac 189
Db 164 ATGATGGGAGCTGCCCGGATGGCGGCTGGCGGACGAAGGACGCCGGTGGAGGACTAC 223

QY 190 gaggcgtgcgcgtgagccactgtccaccacgacatggtgagcgcgcgtggcaggg 249
Db 224 GAGAACTACCGACACGCGCTCGGTGTCCACCACATGACAGCGGAGCGATGGCGCGG 283

QY 250 atcctgagcactgcgtgattgtacccccatgcactgcgtgcacagcccgatgcagagctca 309
Db 284 ATTCTGGAGACTCCATCATGTACCCGGTGGACTCCGCTGGAAGACACGAGATGCAGAGTTG 343

QY 310 cagcctgaccagctgcgcctatcgcaatgtgttgaggccctctgagagattataaga 369
Db 344 AATCCAGATCCCAAGCCCCGGTATACAGCATCTATGGCCGCCCTCAAGAGGATCATGTCAC 403

QY 370 acggaggcctatggagcccatgaggggctgaacgtcacagcaacagcgcgaggcgct 429
Db 404 ACTGAAGGCTTCTGGAGGCCCTCGCGGGCTGGAACGTGATGATGATGGGTGAGGGGCC 463

QY 430 gcccaagcccttatttggctctacgaaaagttaaaaaagacattgagtgtgaatc 489

```


info@bcgsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Jacqueline Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smalish, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 15 Row: h Column: 21
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 7706149.

FEATURES

Location/Qualifiers

1..858
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:8764 IMAGE:3910567"
/tissue_type="Uterus, leiomyosarcoma"
/clone_lib="NIH_MGC_71"
/lab_host="DH10B"
/note="vector: PCMV-SPORT6"
26..493
/codon_start=1
/product="Unknown (protein for MGC:8764)"
/protein_id="AAH15013.1"
/db_xref="GI:15929110"
/translation="MELRSGVSGQAVARRMDGSDRGCKGDKATGSDYENLPTSA
VSTGVNMGAGAILHSEVYDVSVKTRMQSLSPDKAQYTSIYALKIMRKTGFWR
PLTGNNVIMGAGFAHAMFYFACENMKRTLNDVFHHQGNHSLANGILKAFYWS"

CDS

215 a 220 c 231 g 192 t

BASE COUNT

ORIGIN

Query Match

Best Local Similarity 29.6%; Score 172; DB 9; Length 858;

Matches 235; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 183 ggactacaggcgctggcggctggagcactgtccaccagcacatgggtggcaggcgctg 242

Db 127 GGACTACGAGAACCTGGCCACTAGCGCTTCCTGCTGCCACCCACATGACAGCAGGACCGAT 186

QY 243 ggcaggatctctggagcactgctgtgtaccccatcgactcgctcaagaccggatgca 302

Db 187 GCGCGGATCTTGGAGCACTCGTCTATGTACCCGGTGACTCGGTCAAGACACGAATGCA 246

QY 303 gagtctacagcctgaccagctgcccgcgtatcgcaatgtgtggaggccctctggagat 362

Db 247 GAGTTTGTAGTCCAGATGCCAAAGCCAGTACACAAAGTATCTACGGAGCCCTCAAGAAAAAT 306

QY 363 tataaagcaggccctatgagggccatgaggggctgaacgtcacagcaacaggcgc 422

Db 307 CATGCGGACCGAAGGCTCTTGGAGGCGCCTTGGAGCGGCTCAACGTCATGATCATGGGTGC 366

QY 423 agggctgcgcacgcctctttttgctgtcacgaaagttaaaaaagacattgagtga 482

Db 367 AGGGCGGCGCCATGCCTGATTTTGCCTGCTATGAAACATGAAAGGACCTTTAAATGA 426

QY 483 tgtaatccacctgggggcaatagccatttgcctaatggt 522

Db 427 CGTTTTCCACCACCAAGAAACAGCCACCTAGCCCAACGGT 466

RESULT 15

AC099368

LOCUS

DEFINITION Rattus norvegicus clone CH230-84G11, *** SEQUENCING IN PROGRESS

AC099368

AC099368

AC099368.2

AC099368.2

AC099368.2

AC099368.2

AC099368.2

AC099368.2

AC099368.2

AC099368.2

AC099368.2

SOURCE ORGANISM

Norway rat.

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

1 (bases 1 to 163205)

REFERENCE AUTHORS

Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C., Albrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Cogle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Day-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mel,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenko,S., Ogih,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,W., Ruiz,S., Savary,G., Scherer,S., Scott,G., Shen,H., Shoostari,N., Sisson,I., Sodergren,E., Sonaik,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,K., Williams,G., Williamson,A., Wlezyk,R., Wooden,S., Worley,K., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstock,G. and Gibbs,R.

Direct Submission

Unpublished

2 (bases 1 to 163205)

Worley,K.C.

Direct Submission

Submitted (10-NOV-2001)

Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Dec 20, 2001 this sequence version replaced gi:16901901.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GHJ

Center clone name: CH230-84G11

----- Summary Statistics

Assembly program: Phrap; version 0.990329First call to

findPhrapList

Consensus quality: 127058 bases at least Q40

Consensus quality: 138264 bases at least Q30

Consensus quality: 146548 bases at least Q20

Estimated insert size: 120308; sum-of-contigs estimation

Quality coverage: 0x in Q20 bases; agarose-fp estimation

Quality coverage: 1.9x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently
* consists of 71 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 8509: contig of 8509 bp in length
* 8510 8609: gap of unknown length
* 8610 13540: contig of 4931 bp in length
* 13541 13640: gap of unknown length
* 13641 19516: contig of 5876 bp in length
* 19517 19616: gap of unknown length
* 19617 24990: contig of 5374 bp in length
* 24991 25090: gap of unknown length
* 25091 29401: contig of 4311 bp in length
* 29402 29501: gap of unknown length
* 29502 33053: contig of 3552 bp in length
* 33054 33153: gap of unknown length
* 33154 38439: contig of 5286 bp in length
* 38440 38539: gap of unknown length
* 38540 41183: contig of 2644 bp in length
* 41184 41283: gap of unknown length
* 41284 44936: contig of 3653 bp in length
* 44937 45036: gap of unknown length
* 45037 47637: contig of 2601 bp in length
* 47638 47737: gap of unknown length
* 47739 50166: contig of 2429 bp in length
* 50167 50266: gap of unknown length
* 50267 54170: contig of 3904 bp in length
* 54171 54270: gap of unknown length
* 54271 57672: contig of 3402 bp in length
* 57673 57772: gap of unknown length
* 57773 61162: contig of 3390 bp in length
* 61163 61262: gap of unknown length
* 61263 63890: contig of 2628 bp in length
* 63891 63990: gap of unknown length
* 63991 66900: contig of 2910 bp in length
* 66901 67000: gap of unknown length
* 67001 69717: contig of 2717 bp in length
* 69718 69817: gap of unknown length
* 69818 72921: contig of 3104 bp in length
* 72922 73021: gap of unknown length
* 73022 75146: contig of 2125 bp in length
* 75147 75246: gap of unknown length
* 75247 79038: contig of 3792 bp in length
* 79039 79138: gap of unknown length
* 79139 80340: contig of 1202 bp in length
* 80341 80440: gap of unknown length
* 80441 82366: contig of 1926 bp in length
* 82367 82466: gap of unknown length
* 82467 84184: contig of 1718 bp in length
* 84185 84284: gap of unknown length
* 84285 86001: contig of 1717 bp in length
* 86002 86101: gap of unknown length
* 86102 87695: contig of 1594 bp in length
* 87696 87795: gap of unknown length
* 87796 89287: contig of 1492 bp in length
* 89288 89387: gap of unknown length
* 89388 91950: contig of 2563 bp in length
* 91951 92051: gap of unknown length
* 92051 94575: contig of 2525 bp in length
* 94576 94675: gap of unknown length
* 94676 96581: contig of 1906 bp in length
* 96582 96681: gap of unknown length
* 96682 98235: contig of 1554 bp in length
* 98236 98335: gap of unknown length
* 98336 99814: contig of 1479 bp in length
* 99815 99914: gap of unknown length
* 99915 100964: contig of 1050 bp in length
* 101064: gap of unknown length

* 101065 103074: contig of 2010 bp in length
* 103075 103174: gap of unknown length
* 103175 104311: contig of 1137 bp in length
* 104312 104411: gap of unknown length
* 104412 106834: contig of 2423 bp in length
* 106835 106934: gap of unknown length
* 106935 108922: contig of 1988 bp in length
* 108923 108923: gap of unknown length
* 109023 110868: contig of 1846 bp in length
* 110869 110968: gap of unknown length
* 110969 113056: contig of 2088 bp in length
* 113057 113156: gap of unknown length
* 113157 114296: contig of 1140 bp in length
* 114297 114396: gap of unknown length
* 114397 116312: contig of 1916 bp in length
* 116313 116412: gap of unknown length
* 116413 117492: contig of 1080 bp in length
* 117493 117592: gap of unknown length
* 117593 119552: contig of 1960 bp in length
* 119553 119652: gap of unknown length
* 119653 120658: contig of 1006 bp in length
* 120659 120758: gap of unknown length
* 120759 121895: contig of 1137 bp in length
* 121896 121995: gap of unknown length
* 121996 123272: contig of 1277 bp in length
* 123273 123372: gap of unknown length
* 123373 125310: contig of 1938 bp in length
* 125311 125410: gap of unknown length
* 125411 126891: contig of 1481 bp in length
* 126892 126991: gap of unknown length
* 126992 128690: contig of 1699 bp in length
* 128691 128790: gap of unknown length
* 128791 130827: contig of 2037 bp in length
* 130828 130927: gap of unknown length
* 130928 132501: contig of 1574 bp in length
* 132502 132601: gap of unknown length
* 132602 134071: contig of 1470 bp in length
* 134072 134171: gap of unknown length
* 134172 135838: contig of 1667 bp in length
* 135839 135938: gap of unknown length
* 135939 137589: contig of 1651 bp in length
* 137590 137689: gap of unknown length
* 137690 139043: contig of 1354 bp in length
* 139044 139143: gap of unknown length
* 139144 140777: contig of 1634 bp in length
* 140778 140877: gap of unknown length
* 140878 142526: contig of 1649 bp in length
* 142527 142626: gap of unknown length

Query Match 21.5% Score 125.2; DB 2; Length 163205;

Best Local Similarity 87.4%; Pred. No. 2.6e-11;

Matches 194; Conservative 0; Mismatches 23; Indels 5; Gaps 5;

Qy 304 agtctacagcctgaccag-cgtgccgcctatcgcaatgtgttga-ggccctctggagga 361
Db 13720 ATTTTACAGCCTGATCCAGCGGCCGTTATCGAAAAGTTGTGCGAGGCTTTTGGAGAA 13779
Qy 362 ttataagaagcaggcctatggagcccatga-gggggcgtgaacgtcaacagcagc 420
Db 13780 TTATAAGAACAGAGCGCCCTGTGGAGGCCCATGAGGGGGTTGAACCTCCTGCGACTGGC 13839
Qy 421 gcaggcctgccacgccttatttgcctgctacgaaagttaaaaaagacattgagt 480
Db 13840 GCGGGGCTGCCACAGCCCTCTATTTTG-CYGCCTACGAAAAG-TAAAAAGACATTGAGT 13897
Qy 481 gatgtaaccaccctgggggcaatagccattatgccaatggt 522
Db 13898 GATCTAATCCACCCCGGGGCAATAGCCATATTTGCAATGCT 13939

Search completed: September 28, 2002, 07:53:07
Job time: 14508 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model.

Run on: September 28, 2002, 04:53:25 ; Search time 664.31 Seconds
(without alignments)
1504.184 Million cell updates/sec

Title: US-09-870-113-3

Perfect score: 582

Sequence: 1 atgaggttgaggggcgggg.....tgaacctgcggaaggctga 582

Scoring table:

IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_032802.*

```

1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	469.4	80.7	1336	21 AAC90457	Human uncoupling p
2	462	79.4	1418	21 AAC74843	Human ORFX ORF398
3	454.4	78.1	1322	22 AAF27733	Human transport pr
c 4	440	75.6	1294	22 AA160661	Human polynucleoti
5	401	68.9	2502	22 AAS03906	Human secreted pro
6	398.6	68.5	1316	22 AA158875	Human polynucleoti
7	273	45.9	401	22 AAF66156	Novel human polynu
8	233	40.0	27960	22 AAK69779	Human immune/haema
9	233	40.0	27960	22 AAK73320	Human immune/haema

10	200.4	34.4	1902	22	AAK52172	Human polynucleoti
11	196.6	33.8	1305	22	AAF59920	Human mitochondria
12	187.4	32.2	2562	22	AAH90057	Human bone marrow
13	181	31.1	339	22	AAK58672	Human immune/haema
14	174.2	29.9	1890	22	AAK94897	Human full-length
15	173.6	29.8	842	22	AAK92357	Human cDNA 5'-end
16	173.6	29.8	842	22	AAK93895	Human cDNA clone r
17	173.6	29.8	1814	22	AAK94871	Human full-length
18	172.6	29.7	1468	21	AAK77175	Human ORFX ORF2730
19	172	29.6	3028	22	AAK91320	Human polynucleoti
20	170.4	29.3	622	22	AAK91764	Human cDNA 5'-end
21	170.4	29.3	622	22	AAK93200	Human cDNA clone r
22	168.6	29.0	1913	22	AAH89944	Human bone marrow
23	160	27.5	1758	20	AAV84573	Human secreted pro
24	160	27.5	1758	22	ABA83356	Human secreted pro
25	132.8	22.8	1716	21	AAK77189	Human ORFX ORF2744
26	121	20.8	300	20	AAZ13026	Human gene express
27	121	20.8	710	20	AAZ15876	Human gene express
28	109.8	18.9	2837	22	AAK80493	Human immune/haema
29	105.2	18.1	483	24	AB199871	Mouse ischaemic co
30	79.8	13.7	673	21	AAF12614	Aspergillus oryzae
31	72.4	12.4	2137	23	ABL08627	Drosophila melanog
32	67.8	11.6	3058	22	AAK83876	Human immune/haema
33	65.8	11.3	448	22	AAK77174	Human ORFX ORF2729
34	63.2	10.9	1000	21	AAA02484	Human colon cancer
35	59.2	10.2	349	22	AA184504	Human polynucleoti
36	58.2	10.0	318	21	AAA38184	Primer used in the
37	58.2	10.0	320	21	AAA38186	Primer used in the
c 38	58.2	10.0	2188	20	AAZ77506	Human ovarian tumo
c 39	58.2	10.0	3198	20	AAK02974	Human IL-1ra BAC c
40	58	10.0	2037	24	AAK562787	cDNA sequence #574
41	57.6	9.9	320	21	AAA38185	Primer used in the
42	57.6	9.9	1698	23	AAK84408	DNA encoding novel
43	57.6	9.9	2049	23	AAK76210	DNA encoding novel
44	57.4	9.9	320	21	AAA38183	Primer used in the
45	57.2	9.8	1286	24	AB199656	Mouse ischaemic co

ALIGNMENTS

RESULT 1

AAC90457
ID AAC90457 standard; cDNA; 1336 BP.

XX AAC90457;

XX AAC90457;

XX 12-MAR-2001 (first entry)

DT Human uncoupling protein cDNA #6.

DE Human;

DE Human;

DE Human;

DE Human;

DE Human;

DE Human;

DE Human;

DE Human;

DE Human;

DE Human;

DE Human;

DE Human;

DE Human;

DE Human;

DE Human;

DE Human;

DE Human;

DE Human;

DE Human;

DE Human;

DE Human;

DE Human;

DE Human;

DE Human;

DE Human;

DE Human;

DE Human;

DE Human;

DE Human;


```
Db 61 cgggactacgagcgtgcggtgagccactgtcaccacgacacatggtgagcgcgc 120
Qy 241 gtggcaggatcctgagcactgcgtgatgtaccacatgactgcgtcaagaccgcgatg 300
Db 121 gtggcaggatcctgagcactgcgtgatgtaccacatgactgcgtcaagaccgcgatg 180
Qy 301 cagagtctacacgtcagccagctgcgcgtatcgcaatgtgtgagggccctctggagg 360
Db 181 cagagtctacacgtcagccagctgcgcgtatcgcaatgtgtgagggccctctggagg 240
Qy 361 attataaagcaggagcctatggagcccatgaggggctgaaagcgtcacacacagcgc 420
Db 241 attataaagcaggagcctatggagcccatgaggggctgaaagcgtcacacacagcgc 300
Qy 421 gaggggcctgcacacgcctcttatttgcctgctacgaaagttaaaaaagacattgagt 480
Db 301 gcaggggcctgcacacgcctcttatttgcctgctacgaaagttaaaaaagacattgagt 360
Qy 481 gatgtaataccacccctgggggcaatagccatattgccaatgggtgcggcggtgtgtgca 540
Db 361 gatgtaataccacccctgggggcaatagccatattgccaatgggtgcggcggtgtgtgca 420
Qy 541 acattactcatgatgcagccatgaacccctgcggagagctga 582
Db 421 acattactcatgatgcagccatgaacccctgcggagagctga 462

RESULT 3
AAF27733
ID AAF27733 standard; cDNA: 1322 BP.
XX AC
XX AAF27733;
XX
DT 28-MAR-2001 (first entry)
XX
DE Human transport protein TPPT-33 coding sequence.
XX
KW Human; transport protein; TPPT; transport disorder; metabolic disorder;
KW neurological disorder; cardiovascular disorder; reproductive disorder;
KW immune disorder; cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200078953-A2.
XX
XX
PD 28-DEC-2000.
XX
XX
PF 16-JUN-2000; 2000WO-US16668.
XX
PR 17-JUN-1999; 99US-0139923.
PR 10-AUG-1999; 99US-0148177.
PR 18-AUG-1999; 99US-0149357.
PR 28-OCT-1999; 99US-0162287.
XX
FA (INCY-) INCYTE GENOMICS INC.
XX
PI Lal P, Yang J, Yue H, Hillman JL, Tang YT, Bandman O, Burford N;
PI Baughn MR, Azimzai Y, Lu DAM, Au-Young J, Patterson C;
XX
DR WPI; 2001-041424/05.
XX
DR P-PSDB; AAB60113.
XX
XX
Isolated polypeptide with a human transport protein sequence is useful
PT for the diagnosis, prevention and treatment of disorders associated
PT with the immune, reproductive and cardiovascular systems -
XX
PS Claim 5; Page 160; 165pp; English.
XX
XX
The present invention provides the protein and coding sequences for 43
CC novel human transport proteins (designated TPPTs). These can be used in
CC the diagnosis and treatment of transport, metabolic, neurological,
CC reproductive, cardiovascular and immune disorders, and cell proliferative
```

```
CC disorders such as cancer.
XX
SQ Sequence 1322 BP; 328 A; 376 C; 351 G; 267 T; 0 other;
0;

Query Match 78.1%; Score 454.4; DB 22; Length 1322;
Best Local Similarity 98.7%; Pred. NO. 2e-82;
Matches 458; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 114 cggggggccgagcgcgggagggccgctgagggcccggtacgacaaagatcccgga 173
Db 2 cggggggccgagcgcgggagggccgctgagggcccggtacgacaaagatcccgga 61
Qy 174 ctccggcccgactacgagggcgctgcggctggagccactgtcacccgcacatggtggc 233
Db 62 ctccggcccgactacgagggcgctgcggctggagccactgtcacccgcacatggtggc 121
Qy 234 aggcgcctgagggatcctgagcactgctgagccatgcgtatgaccccatcgactgcgtcaagac 293
Db 122 aggcgcctgagggatcctgagcactgctgagccatgcgtatgaccccatcgactgcgtcaagac 181
Qy 294 cgggatgcagagctcacagcctgacccagctgcccgcctatgaggggctgaaagcgtcacagc 353
Db 182 cgggatgcagagctcacagcctgacccagctgcccgcctatgaggggctgaaagcgtcacagc 241
Qy 354 ctggaggattataaagacggagggcctatgagggcccatgaggggctgaaagcgtcacagc 413
Db 242 ctggaggattataaagacggagggcctatgagggcccatgaggggctgaaagcgtcacagc 301
Qy 414 aacaggcgagggcctgcccgccttatttgcctgctacgaaaagttaaaaaagac 473
Db 302 aacaggcgagggcctgcccgccttatttgcctgctacgaaaagttaaaaaagac 361
Qy 474 attgagtgtatgaatccacccctgggggcaatagccatattgccaatggtgcggcggtg 533
Db 362 attgagtgtatgaatccacccctgggggcaatagccatattgccaatggtgcggcggtg 421
Qy 534 tgtggcaacattactcatgatgcagccatgaacccctgcggaag 577
Db 422 tgtggcaacattactcatgatgcagccatgaacccctgcggaag 465

RESULT 4
AAI60661/c
ID AAI60661 standard; cDNA: 1294 BP.
XX AC
XX AAI60661;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 4650.
XX
XX
Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
XX
PD 26-JUL-2001.
XX
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
```


|||||
Db 66 ccggactcggccggactacgagcgctgccgctggagccactgtcaccacgcacatg 125
Qy 229 gfggaggcgcctgagcaggatcctgagagcactcgctgatatcccatcgaactgcgtc 288
Db 126 gfggaggcgcctgagcaggatcctgagagcactcgctgatatcccatcgaactgcgtc 185
Qy 289 aagaccggatcgagatgtacagctgaccagctgcccgtatcgcaaatgtgtggag 348
Db 186 aagaccggatcgagatgtacagctgaccagctgcccgtatcgcaaatgtgtggag 245
Qy 349 gccctctgagatataaagacgagggcctatcgagggcccatgaggggctgaacgtc 408
Db 246 gccctctgagatataaagacgagggcctatcgagggcccatgaggggctgaacgtc 305
Qy 409 acagcaacagcgcgagggcctgcccacgcccctttatttgcctgctacgaaagttaaaa 468
Db 306 acagcaacagcgcgagggcctgcccacgcccctttatttgcctgctacgaaagttaaaa 365
Qy 469 aagacattgagtgatgaatccaccctgggggcaatagccatattgccaatggt 522
Db 366 aagacattgagtgatgaatccaccctgggggcaatagccatattgccaatggt 419
RESULT 6
AAI58875
ID AAI58875 standard; cDNA; 1316 BP.
XX
AC AAI58875;
XX
DF 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 1078.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI: 2001-442253/47.
DR P-PSDB; AAM39719.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Claim 1; SEQ ID NO 1078; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and

the encoded polypeptides (AAM38642-AAM42213) with nootropic,
immunosuppressant and cytostatic activity. The polynucleotides are useful
in gene therapy. A composition containing a polypeptide or polynucleotide
of the invention may be used to treat diseases of the peripheral nervous
system, such as peripheral nervous injuries, peripheral neuropathy and
localised neuropathies and central nervous system diseases, such as
Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
utilisation of the activities such as: Immune system suppression,
Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
and thrombolytic activity, cancer diagnosis and therapy, drug screening,
assays for receptor activity, arthritis and inflammation, leukaemias and
C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 1316 BP; 335 A; 364 C; 337 G; 280 T; 0 other;
Query Match 68.5%; Score 398.6; DB 22; Length 1316;
Best Local Similarity 95.5%; Pred. No. 3.le-71;
Matches 421; Conservative 0; Mismatches 19; Indels 1; Gaps 1;
Qy 138 cggggcctgagggcccggtacgacaaagatccggactccggccggactacgagcgct 197
Db 15 cgaggcctgagggcccggtacgacaaagatccggactccggccggactacgagcgct 74
Qy 198 gccggctgagggcccggtacgacaaagatccggactccggccggactacgagcgct 256
Db 75 gtcggctgagggcccggtacgacaaagatccggactccggccggactacgagcgct 134
Qy 257 agcactgcgtgatgtaccccatcgaactcgtcaagaccggcgatcgagagtcacagcctg 316
Db 135 agcactgcgtgatgtaccccatcgaactcgtcgtgatgaccggcgatcgatgtctacagtcg 194
Qy 317 acccagctgccgctatcgcaatgtgttgaggccctctgagagattataaagaacggagg 376
Db 195 acccagctgccgctatcgcaatgtgttgaggccctctgagagattataaagaacggagg 254
Qy 377 gccatgagggcccggtacgacaaagatccggactccggccggactacgagcgctcccaag 436
Db 255 gccatgagggcccggtacgacaaagatccggactccggccggactacgagcgctcccaag 314
Qy 437 ccccttatttgcctgtacgaaagttataaaagacattgagtgatgaatccaccctg 496
Db 315 ccccttatttgcctgtacgaaagttataaaagacattgagtgatgaatccaccctg 374
Qy 497 ggggcaatagccatattgccaatggcgccgggtgtgtggcaacattacttcattgatg 556
Db 375 ggggcaatagccatattgccaatggcgccgggtgtgtggcaacattacttcattgatg 434
Qy 557 cagccatgaacccctgcggaag 577
Db 435 cagccatgaacccctgcggaag 455
RESULT 7
AAF66156
ID AAF66156 standard; cDNA; 401 BP.
XX
AC AAF66156;
XX
DT 09-APR-2001 (first entry)
XX
DE Novel human polynucleotide, SEQ ID NO: 1912.
XX
KW Human; cytostatic; gene therapy; colon cancer; prostate cancer;
KW breast cancer; lung cancer; cancer detection; ss.
XX
OS Homo sapiens.
XX
PN WO200102568-A2.
XX

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX Homo sapiens.
OS
XX
PN WO200157182-A2.
XX
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251858.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.

```
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
DR
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
PS Disclosure: SEQ ID NO 28132; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 27960 BP; 7998 A; 5835 C; 6583 G; 7544 T; 0 other;

Query Match 40.0%; Score 233; DB 22; Length 27960;
Best Local Similarity 94.2%; Pred. No. 5.6e-38;
Matches 242; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 266 tgatgtacccaatgctgctcaagaccggatgacagatgctacagctgaccagctg 325
DB 18263 tgcgtctctcctgtgtaattcagaccggatgacagatgctacagctgaccagctg 18322
QY 326 ccgcgtatgcgaatgtgttgaggccctctgaggattataaagcaggagccctatga 385
DB 18323 ccgcgtatgcgaatgtgttgaggccctctgaggattataaagcaggagccctatga 18382
QY 386 ggcccatgagggggctgaacgtcacagcaacaggcgagggcctgccacgcccctttatt 445
DB 18383 ggcccatgagggggctgaacgtcacagcaacaggcgagggcctgccacgcccctttatt 18442
QY 446 ttgcctgtacgaaagttaaaagacattgagtgatgataatcacccctgggggcaata 505
DB 18443 ttgcctgtacgaaagttaaaagacattgagtgatgataatcacccctgggggcaata 18502
QY 506 gccattatggcaatggt 522
DB 18503 gccattatggcaatggt 18519

RESULT 10
AAK52172
ID AAK52172 standard; cDNA; 1902 BP.
XX
AC AAK52172;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 717.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
```

```
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
PN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
XX
XX 05-FEB-2001; 2001WO-US04098.
XX
XX 03-FEB-2000; 2000US-0496914.
XX 27-APR-2000; 2000US-0560875.
XX 20-JUN-2000; 2000US-0598075.
XX 19-JUL-2000; 2000US-0620325.
XX 01-SEP-2000; 2000US-0654936.
XX 15-SEP-2000; 2000US-0663561.
XX 20-OCT-2000; 2000US-0693325.
XX 30-NOV-2000; 2000US-0728422.
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
XX Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
XX Xue AJ, Yang Y, Wehrman T, Goodrich R;
XX
XX WPI; 2001-476283/51.
XX P-PSDB; AAM79039.
XX
XX Nucleic acids encoding polypeptides with cytokine-like activities,
XX useful in diagnosis and gene therapy -
XX
XX Claim 1; Page 2450-2452; 6221pp; English.
XX
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
XX encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to
XX cytokine, cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and
XX activin/inhibin activity and may be useful in the diagnosis and/or
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX inflammation.
XX
XX Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
XX (AAM80020) are omitted as the relevant pages from the sequence listing
XX were missing at the time of publication.
XX
XX Sequence 1902 BP; 434 A; 532 C; 547 G; 389 T; 0 other;

Query Match 34.4%; Score 200.4; DB 22; Length 1902;
Best Local Similarity 62.9%; Pred. No. 1.7e-31;
Matches 326; Conservative 0; Mismatches 191; Indels 1; Gaps 1;

QY 61 cggagcccccgggagctggcgtgctggaacgggtgctgacggggcgctggccggggg 120
DB 1040 ctgcgcccccgggagctggcgtggaacgggtgctgacggggcgctggccggggg 1099
QY 121 gccggcgccgggagggccggggcctgcaagcccccgtacgacagatcgcgactccggc 180
DB 1100 gccggcgaggtggtgggacagccgagatggcgcgccggaaggacgacccgggtc 1159
QY 181 -ccggactacgagggcgctgccgctggagccactgtcacccgacatggtggcaggcgc 239
DB 1160 ggaggactacgagaacctgccgactagcgctccgtgtccaccacatgacagcaggagc 1219
QY 240 cgtggcaggagtcctcgtggagcactgctgtagtaccaccatcgactcgctcaagaccggat 299
DB 1220 gatggccggagtcctcgtggagcactgctgtagtaccaccggtggactcgtggaagacgaat 1279
QY 300 gcagagtctacagcctgaccacgctgcccgctatcgcaatgtgttgaggccctctggag 359
```


RESULT	13	
AAK58672		
ID	AAK58672 standard; cDNA; 339 BP.	
XX		
XX	AAK58672;	
XX		
XX	AC	
XX		
XX	06-NOV-2001 (first entry)	
XX		
DE	Human Immune/haematopoietic antigen encoding cDNA SEQ ID NO:3732.	
XX		
XX	Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;	
KW	cystostatic; gene therapy; vaccine; metastasis; ss.	
KW		
XX		
OS	Homo sapiens.	
XX		
PN	WO200157182-A2.	
XX		
PD	09-AUG-2001.	
XX		
XX	17-JAN-2001; 2001WO-US01354.	
PF		
XX		
XX	31-JAN-2000; 2000US-0179065.	
PR	04-FEB-2000; 2000US-0180628.	
PR	24-FEB-2000; 2000US-0184664.	

```

PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249219.
PR 17-NOV-2000; 2000US-0249224.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0256719.
PR 08-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI
XX WPI; 2001-483426/52.
DR
XX P-PSDB; AAM85891.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis.
XX
XX Claim 1; SEQ ID NO 3732; 3071pp + Sequence Listing; English.
PS

```

XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I) amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic activity, and can be used in gene therapy and vaccine production. (I) proteins and polynucleotides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate (I) expression. For example, they may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of (I) by expressing inactive proteins or to supplement the patients own production of (I). Additionally, (I) polynucleotides may be used to produce the secreted (I), by inserting the nucleic acids into a host cell and culturing the cell to express the protein. (I) proteins and polynucleotides may be used to prevent, diagnose and treat immune/haematopoietic-related diseases, especially cancers and cancer metastases of haematopoietic-derived cells. AAK64703 to AAK87694 represent human immune/haematopoietic antigen genomic sequences from the present invention. AAK54942 to AAK54950 and AAM82169 CC represent sequences used in the exemplification of the present invention. XX Sequence 339 BP; 86 A; 84 C; 101 G; 64 T; 4 other;

Query Match 31.1%; Score 181; DB 22; Length 339;
Best Local Similarity 95.1%; Pred. No. 1.2e-27;
Matches 195; Conservative 1; Mismatches 8; Indels 1; Gaps 1;

Qy 291 gaccggatgcagagctacagcctgacccagctgacccgctatcgccgctatcgccgctatggtgagggc 350
|||||
Db 136 gaccggatgcagagctacagcctgacccagctgacccgctatcgccgctatcgccgctatggtgagggc 195
|||||
Qy 351 cctctggaggattataagaacggaggcctatggaggccctatggaggccctatggaggccctatggaggccctac 410
|||||
Db 196 cctctggaggattataagaacggaggcctatggaggccctatggaggccctatggaggccctatggaggccctac 254
|||||
Qy 411 agcaacaggcaggcctgcccacgcccctttatttgcctgctacgaaaagttaaaaaa 470
|||||
Db 255 agcaacaggcaggcctgcccacgcccctttatttgcctgctacgaaaagttaaaaaa 314
|||||
Qy 471 gacattgagtgatgaatccaccct 495
|||||
Db 315 ggcattgagtgatgaatccaccct 339
|||||

RESULT 14
AAK94897
ID AAK94897 standard; cDNA; 1890 BP.
XX
XX AAK94897;
AC
XX
XX 06-NOV-2001 (first entry)
DT
XX Human full-length cDNA, SEQ ID NO: 4109.
XX
XX Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
KW
XX Homo sapiens.
OS
XX EP1130094-A2.
PN
XX
XX 05-SEP-2001.
PD
XX
XX 07-JUL-2000; 2000EP-0114089.
PF
XX
XX 08-JUL-1999; 99JP-0194486.
PR
XX 11-JAN-2000; 2000JP-0118774.
PR
XX 02-MAY-2000; 2000JP-0183765.
PR
XX
XX (HELI-) HELIX RES INST.
PA
XX
XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
XX WPI; 2001-524255/58.
DR

Result No.	Score	Query			DB	ID	Description
		Match	Length				
1	62.4	10.7	319	4	US-09-165-264-8	Sequence 8, Appli	
2	58.2	10.0	318	4	US-09-165-264-12	Sequence 12, Appli	
3	58.2	10.0	320	4	US-09-165-264-14	Sequence 14, Appli	
C 4	58.2	10.0	152331	3	US-09-128-155-16	Sequence 16, Appli	
5	57.8	9.9	320	4	US-09-165-264-7	Sequence 7, Appli	
6	57.6	9.9	320	4	US-09-165-264-13	Sequence 13, Appli	
7	57.4	9.9	320	4	US-09-165-264-11	Sequence 11, Appli	
C 8	50.6	8.7	833	2	US-08-403-852D-3	Sequence 3, Appli	
C 9	50.6	8.7	833	3	US-08-510-646B-3	Sequence 3, Appli	
C 10	50.6	8.7	833	3	US-09-231-818-3	Sequence 3, Appli	
C 11	50.6	8.7	5392	2	US-08-403-852D-1	Sequence 1, Appli	
C 12	50.6	8.7	5392	2	US-08-510-646B-1	Sequence 1, Appli	
C 13	50.6	8.7	5392	4	US-09-231-818-1	Sequence 1, Appli	
14	49.4	8.5	43280	2	US-08-804-277C-1	Sequence 1, Appli	
15	47	8.1	1227	2	US-08-803-800A-1	Sequence 1, Appli	
C 16	46.8	8.0	1931	2	US-09-130-114-2	Sequence 2, Appli	
C 17	46.8	8.0	4257	2	US-08-690-473-1	Sequence 1, Appli	
C 18	46.8	8.0	4257	4	US-09-259-821A-1	Sequence 1, Appli	
C 19	46.8	8.0	4257	4	US-08-843-659-1	Sequence 1, Appli	
20	46.8	8.0	12001	1	US-08-458-568A-11	Sequence 11, Appli	
21	46.6	8.0	4403765	4	US-09-103-840A-2	Sequence 2, Appli	
22	46.6	8.0	4411529	4	US-09-103-840A-1	Sequence 1, Appli	
C 23	46.2	7.9	2823	2	US-08-398-008A-1	Sequence 1, Appli	
C 24	46.2	7.9	2823	1	US-08-893-333-1	Sequence 1, Appli	
25	46.2	7.9	3415	1	US-08-054-077C-1	Sequence 1, Appli	
26	45.8	7.9	460	2	US-08-903-800A-4	Sequence 4, Appli	
C 27	45.6	7.8	1300	4	US-08-483-533-39	Sequence 39, Appli	


```

; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: S.pristinaespiralis
; US-09-231-818-1

Query Match      8.7%   Score 50.6;   DB 4;   Length 5392;
Best Local Similarity 53.0%; Pred.No.0.0058;
Matches 133; Conservative 0; Mismatches 114; Indels 4; Gaps 1;

QY 3 ggaattgagggcggggtgctgctgctgctgagcggtgctgacgagggcgccggcgccgagggcccgggcg 62
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3478 GGCCTTCGCGCAGGGTGTGTCGCCGCGTGCCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3419

QY 63 gagcc----ccggggagtcgagcgctgctgagcggtgctgacgagggcgctgagggccgggg 118
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3418 CGCGCAGGACCGGGTCTCGGAGCGCGGTGAGCGCCACCGCTGAAGCCGCTCCGCGGGGCTGT 3359

QY 119 gggccgcgcgggcgggagcgcgggcctgcagggcccggtacgacgaagatccggaactccg 178
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3358 CGCGCGGCAGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 3299

QY 179 gcccgagactcagagcgctgcccggctggagccactgtcaccacgcacatggtggcagcg 238
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3298 GCTCTTCACGAGCGCGCTCGCGCGCGCGCTCGCGCGCGCGCTCGAGCAGGACGAGGTGCGCGCGCGG 3239

QY 239 ccgtggcgagg 249
   ||| ||| |||
Db 3238 CGAGGGCGGGG 3228

RESULT 14
US-08-804-227C-1
; Sequence 1, Application US/08804227C
; Patent No. 5876991
; GENERAL INFORMATION:
; APPLICANT: DeHoff, Bradley S.
; APPLICANT: Kuhstoss, Stuart A.
; APPLICANT: Rosteck, Paul R., Jr.
; APPLICANT: Sutton, Kimberly L.
; TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: THOMAS G. PLANT 1501
; STREET: LILLY CORPORATE CENTER
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: ASCII(DOS) Text only
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,227C
; FILING DATE: February 21, 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Plant, Thomas, G.
; REGISTRATION NUMBER: 35,784
; REFERENCE/DOCKET NUMBER: X-8231
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-2459
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 43280 base pairs
; TYPE: nucleic acid

```


THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2002, 03:49:39 ; Search time 4881.77 Seconds
(without alignments)
1609.094 Million cell updates/sec

Title: us-09-870-113-3
Perfect score: 582
Sequence: 1 atggagttgagggcgggg.....tgaacctgcggaaggtga 582

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	532.6	91.5	888	9	AL530804
2	450.6	77.4	817	10	BI103329
3	438.8	75.4	1131	10	BG295496
4	420.8	72.3	573	10	BI541703
5	409.8	70.4	697	10	BI854643
6	398.2	68.4	967	10	BI739710
7	387	66.5	645	10	BG964218
8	367.2	63.1	397	9	BE012485
9	354.6	60.9	574	10	BI185580
10	346.4	59.5	681	9	AV704087
11	345	59.3	598	10	BE913718
12	323.4	55.6	560	10	BI794616
13	320.2	55.0	480	10	BI045863
14	302.2	51.9	336	9	AW326482
15	298.4	51.3	546	10	BM488747
16	292	50.2	546	9	BE014800
17	290.8	50.0	900	10	BG325430

18	288.6	49.6	747	10	BI827322
19	282.8	48.6	501	9	AW211366
20	274.4	47.1	916	10	BI751780
21	255.8	44.0	622	10	BJ039504
22	244.6	42.0	722	10	BI876051
23	243.6	41.9	417	9	AA104365
24	240.8	41.4	855	10	BI412175
25	240.2	41.3	261	10	BI338246
26	231.4	39.8	701	10	BI831943
27	222.6	38.2	421	10	BI448355
28	207.4	35.6	678	9	BB654057
29	203.2	34.9	561	9	AI942584
30	196.6	33.8	684	9	AI133696
31	196.6	33.8	750	10	BI831750
32	194.6	33.4	729	10	BI460337
33	194	33.3	567	10	BJ035232
34	193.6	33.3	512	9	AW210205
35	191.6	32.9	453	10	BF916224
36	190.2	32.7	612	10	BJ012200
37	187.4	32.2	1295	11	AK015770
38	186.8	32.1	803	10	BG777404
39	184.6	31.7	217	9	AW346915
40	182.8	31.4	331	10	BI449976
41	181.6	31.2	1209	11	AK006155
42	177.8	30.5	666	9	BB644566
43	177.2	30.4	513	10	BI466991
44	174.2	29.9	866	10	BF797743
45	172	29.6	592	10	BM313240

ALIGNMENTS

RESULT 1

AL530804

LOCUS

DEFINITION

AL530804

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

AL530804 AL530804 LTI_NFL001_NBC4 Homo sapiens cdna linear EST 13-FEB-2001
prime, mRNA sequence.

AL530804

AL530804.1 GI:12794297

EST.

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 888)

Li.W.B., Gruber,C., Jessee,J. and Polayes,D.

Full-length cdna libraries and normalization

Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

Location/Qualifiers

1..888

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="CSDD008YH17"

/sex="male"

/tissue_type="neuroblastoma cells"

/lab_host="DH10B"

/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cdna

was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cdna was digested with Not I and

cloned into the Not I and Eco RV sites of the pCMVSPORT 6

vector. Library was normalized. Library was constructed

by Life Technologies. Contact : Peng Liang Life

Technologies, a division of Invitrogen 9800 Medical Center

Drive Rockville, Maryland 20850, USA Fax : (1) 301 610

8371 Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com"

BASE COUNT

168 a

244 c

322 g

139 t

15 others


```
QY 528 cgggtgtgtggaacattacttcattgatgcagccatgaacacctgcggaag 577
||||| || ||||||| ||||||| ||||||| || || |||||
Db 365 CGGGTGGT-ANAACTATTCTTATGATGACGCCATGAATCCAGCAGAAG 413

RESULT 10
AV704087 AV704087 681 bp mRNA linear EST 09-OCT-2000
LOCUS AV704087 ADB Homo sapiens cDNA clone ADBAF05 5', mRNA sequence.
DEFINITION AV704087
ACCESSION AV704087
VERSION AV704087.1 GI:10721407
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 681)
AUTHORS Peng, Y., Song, H., Huang, Q., Huang, C., Gu, Y., Yang, Y., Gao, G., Xiao
, H., Xu, X., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z.,
Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu
, G., Hu, R., Chen, J., Chen, Z., and Han, Z.
TITLE Homo sapiens cDNA ADB clones
JOURNAL Unpublished (2000)
COMMENT Contact: Zequang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shouying Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzq@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
source
Location/Qualifiers
1..681
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="ADBAFE05"
/clone_lib="ADB"
/tissue_type="Adrenal gland"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 149 a 209 c 205 g 117 t 1 others
ORIGIN

Query Match 59.5%; Score 346.4; DB 9; Length 681;
Best Local Similarity 99.7%; Pred. No. 8.1e-54;
Matches 347; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 116 gggggccggcgccggggagcgccgggctgcagggcccggtacgacaagatccggact 175
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1 GGGGGCCGGCGCGGGAGGCGCGGGCTGCGAGGCCCGCGGTACGACAAATCCGGACT 60

QY 176 cggggccggactacagggcgctgcgggtggagccactgtcaccccgacatggtggcag 235
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 61 CGGGCCGGGACTACAGAGCGCTGCCGGCTGGAGCCACTGTCAACCGACATGTTGGCAG 120

QY 236 gcgcctgtggcaggatcctggagcactcgtgtatgccctcactgcgtcgaagacc 295
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 121 GGGCCGAGCGAGGATCTGAGGCACTCGGTGATGTACCCCATCGACTCGGTCAAGACCC 180

QY 296 ggatcagagtacagcctgacccagctgcccgctatcgcaatgtgtggaggccctct 355
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 181 GGATCGAGAGTACAGGCTGACCCAGCTGCCGCTATCGCAATGTGTGGAGGCCCTCT 240

QY 356 ggagattataagaacgaggccctatggagcccatagggggcgtacgctcacagcaa 415
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 241 GGAGGATTATAAGAACGAGGGCCCTATGAGGCCCATGAGGGGCTGAACCTCAGACAA 300

QY 416 caggcgaggccctgccacgccctttatttgcctgtacgaaaaagt 463
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 301 CAGGCGCAGGGCCCTGCCACGCCCTTTATTTCCTGTCTACGAAAAGT 348
```

```
RESULT 11
BE913718 BE913718 598 bp mRNA linear EST 29-SEP-2000
LOCUS BE913718
DEFINITION BE913718
ACCESSION BE913718
VERSION BE913718.1 GI:10411618
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 598)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9146 row: o column: 17
High quality sequence stop: 593.
Location/Qualifiers
1..598
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:3969256"
/clone_lib="NCI_CGAP_Maml"
/tissue_type="tumor, biopsy sample"
/dev_stage="10 months, virgin"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

BASE COUNT 141 a 175 c 172 g 110 t
ORIGIN
```

```
Query Match 59.3%; Score 345; DB 10; Length 598;
Best Local Similarity 93.5%; Pred. No. 1.5e-53;
Matches 360; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 195 gctgcgctgagccactgtcacccacgacatgtggcagcgccgtggcaggatcct 254
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1 GCTGCCGCTGAGGCCACTGTCAACCCACACATGTGGCGCCCGTGGCAGGATCCT 60

QY 255 ggagcactgcgtgatgtaccccatcactgcgtcgaagcccgatgcagagtctacagcc 314
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 61 GGAGCATTCGCTGATGTACCCGATCGACTGCGTCAAGACCCGGATGCAGAGCTACAGCC 120

QY 315 tgaccacgtcccgctatcgcaatgtttggaggccctctggaggaattataagaacgga 374
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 121 TGACCCACCGCCCGCTATCGGAACGTTGTGAGAGCTCTCTGAGAAATCATGAGACAGA 180

QY 375 gggccatggaggcccatgagggtgaagtcacacagcgcagggcctgccca 434
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 181 GGGCCTGTGGAGGCCATCGCGGGGCTGAACGTACACACACAGCGCGGGCCTGCCCA 240

QY 435 cgcccttattttgcctgtcacgaaagttaaaagacattgagtgaataccacc 494
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 241 CGCCCTCTATTTCCTGCTACGAAAAGTTAAAAAGACATTGAGTACGTAATCCACC 300

QY 495 tgggggcaatgacccatattgccaattgctgcggccgggtgtgtggaacattactcatga 554
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
```

Db 301 AGGGGCAATAGCCATATGTCATGTCAGCCGGATGTGTGGCAGCATTTACTTCATGA 360

Qy 555 tcagccatgaacctgcggaagcc 579
 |||||
 Db 361 TGCAGCCATGAATCCAGCGGAGTC 395

RESULT 12
 B1794616 560 bp mRNA linear EST 01-OCT-2001
 LOCUS ic92e02.y1 Melton Normalized Mixed Mouse Pancreas 1 NI-MMS1 Mus
 DEFINITION musculus cDNA 5' similar to TR:Q23125 Q23125 W02B12.9 PROTEIN. [1]
 ; mRNA sequence.

ACCESSION B1794616
 VERSION B1794616.1 GI:15822341
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 560)
 Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,
 Lemishka, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
 Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,
 Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas
 , M., Gibbons, M., McCann, R., Cole, R., Tsagareishvili, R., Williams, T.
 , Jackson, Y. and Bowers, Y.
 Endocrine Pancreas Consortium
 Unpublished (2000)
 Other_ESTs: ic92e02.x1
 Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
 Endocrine Pancreas Consortium
 Harvard University, Howard Hughes Medical Institute
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
 MA 02138
 Tel: 617-495-1812
 Fax: 617-495-8557
 Email: dmelton@biohp.harvard.edu
 Library was constructed by Dr. Douglas Melton DNA sequencing by:
 Washington University Genome Sequencing Center For information on
 obtaining a clone please contact: Juliana Brown
 (brown@fas.harvard.edu)
 Seq primer: -40RP from Gibco
 High quality sequence stop: 430.

FEATURES
 source
 1..560
 /organism="Mus musculus"
 /strain="ICR"
 /db_xref="taxon:10090"
 /clone_lib="Melton Normalized Mixed Mouse Pancreas 1
 NI-MMS1"
 /sex="Both for embryonic & newborn, male for adult and
 adult islet"
 /dev_stage="Embryonic day 10.5, E12.5, E16.5, newborn,
 adult, mixed"
 /lab_host="DH10B"
 /notes="Vector: pSPORT1; Site.1: Not I; Site.2: Sal I; Five
 libraries representing E10.5/12.5 pancreatic bud, E16.5
 pancreas, newborn pancreas, adult pancreas, and adult
 islets of Langerhans were separately constructed using
 SuperScript plasmid Library kit (Life Technologies). cDNA
 was made by oligo-dT priming and size-selected by column
 fractionation. Libraries were amplified once on solid
 support and plasmid DNA from each library was prepared
 and mixed in equal amounts. The mixed library DNA was
 normalized by method #4 from Bonaldo, Lennon, and Soares
 1996 Genome Research 6:791-806; 0.5 microgram
 single-stranded mixed library plasmid DNA was mixed with
 5 micrograms PCR product representing mixed library
 inserts and hybridized to an EcoT of 6. Single-stranded
 (unhybridized) plasmids were isolated by hydroxyapatite
 chromatography and used to make this library."

BASE COUNT 134 a 158 c 163 g 105 t

ORIGIN

Query Match 55.6%; Score 323.4; DB 10; Length 560;
 Best Local Similarity 92.9%; Pred. No. 1.3e-49;
 Matches 339; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy 215 tcaccacgacatgtgtgcagcgccctgacaggtatcctgagcactgctgatgtacc 274
 |||||
 Db 1 TCACCACGCACATGTCGGGGCCCGCTGGCAGGATCCTGGAGCATTCGCTGATGTACC 60
 Qy 275 ccattcgactgctgaacagcccggtgagagctctacagctgacccagctgcgcgcctatc 334
 |||||
 Db 61 CGATCGACTGCTCAAGACCCGGATGCAGAGCCTACAGCCTGACCCAGCCGCCCTATC 120
 Qy 335 gcaatgtgtgagggccctctgagagattataagaacgagggcctatgagggccatga 394
 |||||
 Db 121 GGAACGTGTTGGAGGCTCTCTGGAGAATCATGAGGACAGAGGGCTGTGGAGGCCCATGC 180
 Qy 395 gggggtgaaagtcacagcaagcgagcgacctgccacgccccttatttttcctgct 454
 |||||
 Db 181 GGGGCTGAACGTCACAGTACAGCGCGGGCCCTGCCACGCCCTCTATTTTGCCTGCT 240
 Qy 455 acgaaaagttaaaaaagacattgagtgatgtaattccacctgggggcaatagccatttg 514
 |||||
 Db 241 ACGAAAAGTTAAAAAGACATTTAGTGACGTAAATCCACCCAGGGGCAATAGCCATATTG 300
 Qy 515 ccaatgtgagggcggtgtgtgcaacattacttcatgatgacccatgaacccctgagg 574
 |||||
 Db 301 CCAATGTGCGCGGATGTGTGGCAGCATTTACTTCATGATGCGCATTAATCCAGCGG 360
 Qy 575 aagcg 579
 |||||
 Db 361 AAGTC 365

RESULT 13

B1045863
 LOCUS B1045863 480 bp mRNA linear EST 14-JUN-2001
 DEFINITION MR3-FN0209-300101-004-h10 FN0209 Homo sapiens cDNA, mRNA sequence.
 ACCESSION B1045863
 VERSION B1045863.1 GI:14452485
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 480)
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
 Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
 , M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 20202663
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR3&st2=MR3-FN0209-
 300101-004-h10&st3=2001-01-30&st4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 480.
 Location/Qualifiers

FEATURES

```
source
1..480
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="FN0209"
/dev_stage="Adult"
/note="Organ: prostate_normal; Vector: puc18; Site_1: SmaI
; Site_2: SmaI; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. Letters Patent
application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the puc 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."
BASE COUNT      116 a   126 c   146 g   92 t
ORIGIN

Query Match      55.0%; Score 320.2; DB 10; Length 480;
Best Local Similarity 96.0%; Pred. No. 4.9e-49;
Matches 339; Conservative 0; Mismatches 13; Indels 1; Gaps 1;

Qy 226 atggtggcaggcgctggcaggatctctggagcactgcgtgatgacccatcgactgc 285
      |||||
Db 2 AGGTGACACGCGATGTGCGAGGATGTGGAGCGACTGCGTGATACCCATCGACTGC 61

Qy 286 gtcaagaccggatgcagagtctacagcctgaccagctgcccgctatcgcaatgtgtg 345
      |||||
Db 62 GTCAAGACCGGATGCAGAGTCTACAGCTGACCCAGCTGCCGCTATCGCAATGTGTG 121

Qy 346 gaggcccttgagattataaagcaggaggccctatagaggccatgaggggctgaac 405
      |||||
Db 122 GAGGCCCTCTGAGGATTATAGAACGAGGAGGCCCTATGGAGGCCCATGAGGGGCTGAAC 181

Qy 406 gtcaagcaaacaggcgaggcgctgccacgccctttatttgctgctagcaaaagtta 465
      |||||
Db 182 GTCACAGCAACAGGCGAGGGCTGCCACCGCGCTTATTGTCCTGTACGAAAGTTA 241

Qy 466 aaaaagacatgagtgtatgtaataccacct-999ggcaatagccaatattgccaatggtgc 524
      |||||
Db 242 AAAAAAGACATTGAGTGTGTATCCACCCTCTGGGGGGCAATAGCCATATTGCGCAATGTTGC 301

Qy 525 ggcgggtgtgtggcaacttactctatgtagcagccatgaacctgcggaag 577
      |||||
Db 302 GGCCGGGTGTGGCACCATTACTTCTCATGTGACCCATGAACCTCGCGGAG 354

RESULT 14
LOCUS      AW326482      336 bp      mRNA      linear      EST 25-APR-2001
DEFINITION 19106 MARC 1B0V Bos taurus cDNA 5', mRNA sequence.
ACCESSION  AW326482
VERSION     AW326482.1 GI:6762403
KEYWORDS   EST.
SOURCE     COW.
ORGANISM   Bos taurus
REFERENCE  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
          Bovidae; Bovinae; Bos.
AUTHORS   Smith,T.P.L., Grosse,W.M., Ereking,B.A., Roberts,A.J., Stone,R.T.,
          Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett
          ,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-Mckown,C.G.,
          Pertea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and
          Keele,J.W.
TITLE      Sequence evaluation of four pooled-tissue normalized bovine cDNA
          libraries and construction of a gene index for cattle
JOURNAL   Genome Res. 11 (4), 626-630 (2001)
MEDLINE   21180013
COMMENT   Contact: Smith TPL
          USDA, ARS, US Meat Animal Research Center
          PO Box 166, Clay Center, NE 68933-0166, USA
          Tel: 402 762 4366
          Fax: 402 762 4390
          Email: smith@mail.marc.usda.gov

Single pass sequencing. Bases called and trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 20
and -minmatch 12 options.
PCR Primers
FORWARD: AGGAAACAGCTATGACCAT
BACKWARD: GTTTCACAGTCACGACG
Plate: 8 row: F column: 16
Seq primer: ATTTAGGTGACACTATAG.
FEATURES
Location/Qualifiers
source      1..336
            /organism="Bos taurus"
            /db_xref="taxon:9913"
            /clone_lib="MARC 1B0V"
            /tissue_type="pooled"
            /lab_host="DH10B"
            /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
            Library made from pooled tissue from lymph node, ovary,
            fat, hypothalamus, and pituitary."
BASE COUNT      72 a   96 c   107 g   61 t
ORIGIN

Query Match      51.9%; Score 302.2; DB 9; Length 336;
Best Local Similarity 96.0%; Pred. No. 9.4e-46;
Matches 310; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 216 caccagcacatgtgtggcaggcgctggcaggatcctgagcactgcgtgatacc 275
      |||||
Db 14 CACCACGCACATGTGTGGCGGCGCGTGGCAGGATCTCGGACACTCGGTGATGATACC 73

Qy 276 catgactgcgtcaagacccggatgtagagctctacagctgacccagctgccgctatcg 335
      |||||
Db 74 CGTCGACTCGCTCAAGACCCGGATGTCAGAGCTACAGCGGATCCAGCGCCGCTATCG 133

Qy 336 caatgtgtgaggccctctgaggagattataagaacgaggagccctatgaggccatgag 395
      |||||
Db 134 CAATGTGTGTGAGGCCCTCTGGAGGATTTAAGAACGAGGCGCTGTGGAGGCCCATGCG 193

Qy 396 ggggctgaactgcagcaacagcgaggcgctgcccacgcccttatttgcctgcta 455
      |||||
Db 194 AGGCTGAACGCTCAGCAACAGCGGAGGCGCTGCCACGCCCTCTATTGCTGCTGCTA 253

Qy 456 cgaagaagttaaaagacattgagtgtatgtaatccacctgggggcaatagccatttgc 515
      |||||
Db 254 TCAAAAGTTAAAAAGACATTGAGTGTGATTAATCACCTGGGGCAATAGCCATATTGC 313

Qy 516 caatggtgcggcggtgtgtgg 538
      |||||
Db 314 CAATGGTGCGGCTGGGTGTGTGG 336

RESULT 15
LOCUS      BM488747      546 bp      mRNA      linear      EST 07-FEB-2002
DEFINITION pgm2n.pk008.l2 Normalized Chicken Breast Muscle, Leg Muscle, and
          Epiphyseal Growth Plate cDNA library (pgm2n) Gallus gallus cDNA
          clone pgm2n.pk008.l2 5', similar to ref|XP_050766.1 (XM_050766)
          putative mitochondrial solute carrier [Homo sapiens]
          gb|AAK49519.1|AF327402.1 (AF327402) putative mitochondrial solute
          carrier splice variant [Homo sapiens], mRNA sequence.
ACCESSION  BM488747
VERSION     BM488747
KEYWORDS   EST.
SOURCE     chicken.
ORGANISM   Gallus gallus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
          Phasianinae; Gallus.
REFERENCE  1 (bases 1 to 546)
          Cogburn,L.A. and Monsonego-Ornan,E.
          ESTs from Normalized Chicken Breast Muscle, Leg Muscle, and
          Epiphyseal Growth Plate cDNA library, USDA/IFAFS Animal Genome
          Project
```

Unpublished (2002)
Contact: Larry A. Coghurn
University of Delaware
Townsend Hall, Newark, DE 19717, USA
Tel: 302-831-1335
Fax: 302-831-2822
Email: coghurn@udel.edu, www.chickies.com

FEATURES

```

1. .546
/organism="Gallus gallus"
/strain="Commercial broiler and Ottawa Res. Centre
Strains 90 & 21"
/db_xref="taxon:9031"
/clone_pgm2="pk008.12"
/clone_lib="Normalized Chicken Breast Muscle, Leg Muscle,
and Epiphyseal Growth Plate cDNA library (pgm2n)"
/sex="Male and Female"
/tissue_type="Breast muscle, leg muscle and epiphyseal
growth plate"
dev_stage="Breast; leg: Embryo(d19); post-hatch(1d,1,3,5,7,9,
11 weeks); growth plate(1d,7d,14d post-hatch)"
/lab_host="E. coli EMBO10B"
/notes="vector: pCMVSPORT6; Library made from equivalent
pools of total RNA isolated from each tissue (embryonic
muscle 33.3%; juvenile muscle 33.3%; and epiphyseal growth
plate 33.3% of the final RNA pool). Single pass sequencing
from 5'-end"
a 183 c 197 g 70 t 6 others

```

BASE COUNT
ORIGIN

Query Match	51.3%	Score	298.4	DB	10	Length	546
Best Local Similarity	77.0%	Pred. No.	4.6e-45				
Matches	375	Conservative	0	Mismatches	109	Indels	3
						Gaps	1
QY	91	gggtgctgcagcggcgctggcgcggggccgcgcgcggcgagggagggcgagcgtgcagg	150				
Db	13	GGCCGGCGCGCGGGCGCGCGAGCGTCTGGGCCCCCGGGGAGCTGTAGCGGGCTCGG	72				
QY	151	ccccgcgtacgacaagatccgcgactccgcgcgcgactacgagggcgtccgcgtggagcc	210				
Db	73	GGCCCGGAGCGCGCTCGCGC--CCGCCCCGACTACGAGGCGCTGCGCAGGGCGCC	129				
QY	211	actgtccaccagcacatggtgcaggcgcctggcaggatcctcggagcactgcgtgatg	270				
Db	130	GCCGTGTCCAGCAGCATGCTGTGGCGGGCGCCGTGGCGGCATCATGGAGCATGCGCTGATG	189				
QY	271	taccctatcactcgtccaagaccgcgatcagagttctacagctgcacccagctgccgcg	330				
Db	190	TACCCCTCGATTGCTCAAGACTCTGGATGTCAGAGCTTGGCGCCGAGCCCGCCCGCCGC	249				
QY	331	tatcgcaatgtgtggaggccctctggaggattataagaacggaggcgtatggagccc	390				
Db	250	TACCGCAAGCTGTGGAGGCCCTGTGGCGCATCTGTCGTACCGAGGGCGTGTGAGGGCCC	309				
QY	391	atgaggggcgtgaacgtccacgaacaggcgagggcctgcaccaagcccttattttgcc	450				
Db	310	ATGCGGGCGCTGAACATCATCCGACCGCGCGGGCGCGGCCACCGCCCTCTACTTCGCC	369				
QY	451	tgctacaaaagttaaaaagacattgaigtatgaattccaccctgggggcaatagccat	510				
Db	370	TGCTACGAAAAGTTTAAAAAGACGCTGAGCGACGTCATCCACGCGGGGGGCAATAGCCAT	429				
QY	511	attgccaatggtgcgcgggtgtgtggcaacttacttcatgatgatgagccatgaacct	570				
Db	430	GTGCNNNGGTGCAGCGGGTGTGTAGCAACATTGCTCCACGACGACGATGACCCCT	489				
QY	571	gcggaag	577				
Db	490	GCAGAAG	496				

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2002, 07:53:07 ; Search time 5287.54 Seconds
(without alignments)
2742.691 Million cell updates/sec

Title: US-09-870-113-5

Perfect score: 693

Sequence: 1 atggagttgaggcgggg.....ttacctgtttaccactag 693

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10453268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_on.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_on.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htgo_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description

1	577	83.3	1448	9	AF327402	AF327402 Homo sapi
2	577	83.3	1448	17	HSA303077	AJ303077 Homo sapi
3	520.4	75.1	1889	9	AF327403	AF327403 Homo sapi
4	520.4	75.1	1889	17	HSA303078	AJ303078 Homo sapi
5	454.4	65.6	1322	6	AX061229	AX061229 Sequence
6	291	42.0	123160	9	AL353719	AL353719 Human DNA
7	273	39.4	401	6	AX071440	AX071440 Sequence
8	251	36.2	221062	2	AC096351	AC096351 Rattus no
9	207.6	30.4	221062	2	AC096351	AC096351 Rattus no
10	207.2	29.9	4112	10	AF288621	AF288621 Mus muscu
11	198.8	28.7	1429	9	AF223466	AF223466 Homo sapi
12	177.8	25.7	5869	10	AF361699	AF361699 Mus muscu
13	174.2	25.1	3415	9	AY032628	AY032628 Homo sapi
14	172	24.8	858	9	BC015013	BC015013 Homo sapi
15	125.2	18.1	163205	2	AC099368	AC099368 Rattus no
16	116.2	16.8	191702	2	AC108878	AC108878 Mus muscu
17	113.8	16.4	108765	2	AF216674	AF216674 Homo sapi
18	112.2	16.2	174445	9	AC051642	AC051642 Homo sapi
19	112.2	16.2	182504	2	AC022597	AC022597 Homo sapi
20	104.8	15.1	483	6	AX306258	AX306258 Sequence
21	82.4	11.9	34305	3	CEW02B12	Z66521 Caenorhabdi
22	72.4	10.4	3037	3	AY060268	AY060268 Drosophil
23	71.6	10.3	146743	2	AP003518	AP003518 Oryza sat
24	70.8	10.2	2878	3	AF217402	AF217402 Drosophil
25	68	9.8	615	11	CNS06FGL	AL396379 T7 end of
26	68	9.8	71053	2	AC102327	AC102327 Mus muscu
27	67.2	9.7	77191	2	AC023381	AC023381 Homo sapi
28	67.2	9.7	207420	2	AC078884	AC078884 Mus muscu
29	67	9.7	121146	2	AC108558	AC108558 Rattus no
30	66.6	9.6	38723	2	AC095142	AC095142 Rattus no
31	66.6	9.6	59632	2	AC102213	AC102213 Mus muscu
32	66.6	9.6	94914	2	AC105592	AC105592 Rattus no
33	66.2	9.6	158167	2	AC034105	AC034105 Homo sapi
34	66	9.5	65928	2	AC012570	AC012570 Homo sapi
35	65.8	9.5	79355	2	AC021280	AC021280 Homo sapi
36	65.6	9.5	46181	2	AC105855	AC105855 Rattus no
37	65.6	9.5	115715	2	AC079411	AC079411 Homo sapi
38	65.6	9.5	140469	2	AC025066	AC025066 Homo sapi
39	65.6	9.5	148407	2	AC106354	AC106354 Rattus no
40	65.2	9.4	1560	14	HS2IE	M29384 Herpes simp
41	65.2	9.4	62649	2	AC022552	AC022552 Homo sapi
42	65.2	9.4	146744	2	AC098461	AC098461 Rattus no
43	65.2	9.4	174096	2	AC097177	AC097177 Rattus no
44	65	9.4	93663	2	AC107039	AC107039 Pan trogl
45	64.8	9.4	60984	2	AC102552	AC102552 Mus muscu

ALIGNMENTS

RESULT 1

AF327402

LOCUS

DEFINITION

AF327402

ACCESION

VERSION

KEYWORDS

SOURCE

ORGANISM

human.

human sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

1 (bases 1 to 1448)

Li.F.Y., Nikali.K., Grogan,J., Leibiger,I., Leibiger,B.,

Schweyen,R., Larsson,C. and Suomalainen,A.

Characterization of a novel human putative mitochondrial

transporter homologous to the yeast mitochondrial RNA splicing

proteins 3 and 4

FEBIS Lett. 494 (1-2), 79-84 (2001)

21195335

PUBMED

11297739

REFERENCE

2 (bases 1 to 1448)

Nikali K., Human Molecular Genetics, National Public Health Institute,
Mannerheimintie 166, Helsinki, FI-00300, FINLAND.

[3]
Li F., Nikali K., Gregan J., Leibiger I., Leibiger B., Schweyen R.,
Larsson C., Suomalainen A.;
"Characterization of a novel human putative mitochondrial transporter
homologous to the yeast mitochondrial RNA splicing proteins 3 and 4";
FEBS Lett. 494:79-84(2001).

Key Location/Qualifiers

source 1..1889
/db_xref="taxon:9606"
/organism="Homo sapiens"
1032..1565
/gene="HMR3/4"
/product="mitochondrial RNA splicing protein 3/4"
/function="putative mitochondrial solute carrier"
/protein_id="CAC27997.1"
/translation="MNPAEVVKRMQWYNSPHYRVTDCRVWNEGAGAFYRSYTTOL
TMNPFQAIHMFYEFLOHFNPRYNPSHVLSCACAGAAATPLDCKTLLNT
QESLAINSHITGHITMASAFRYVGVGTAFRGVQARVIQIPSTALANSVYEFFK
YLITKQEWRAK"

Sequence 1889 BP: 404 A; 508 C; 544 G; 433 T; 0 other;

Query Match 75.1%; Score 520.4; DB 17; Length 1889;
Best Local Similarity 99.8%; Pred. No. 9.4e-79;
Matches 521; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Y 1 atgaggttgagggcggggtgctggcggtggtggcgggggcgcgcgagggccggg 60
D 9 ATGAGGTGGAGGGCGGGGTGCTGGCGGTGTGCGGGGGGGCAGCGGCGCGGG 68
Y 61 cggagccccgggagtcggcgctgctggaacgggtggtgctgagcggggctggtggcgggg 120
D 69 CGGAGCCCGGGAGTGGCGCTGCTGGAGCGGTGGCTGCAGCGGGGCGTGGCGGGGG 128
Y 121 gccgagcgaggagggcgggcgctgagggcgcccgccggtgagcaagatccggactccggc 180
D 129 GCCGGCGGGGAGGCGCGGGCGCTGCAGGCCCCCGGTACGACAAATCCGGACTCCGGC 188
Y 181 ccggactacagggcgctggcgctggagccactgtcaccacgacatggtggcagggcgc 240
D 189 CCGGACTACGAGGGCGCTGCCGGCTGGAGCCACTGTCTACCGACGACATGTTGGCAGCGCC 248
Y 241 gtggcaggatccctggagcactgctgtagtaccatcgactcgactgctcaagaccggatg 300
D 249 GTGGCAGGGATCCCTGGAGCACTGGCGTGTATGTACCCCATCGACTGCGTCAAGACCGGATG 308
Y 301 cagagttacagctgacccagctgcccgtatgctgaatgtgttggaggccctctggagg 360
D 309 CAGAGTCTACAGCTGACCCAGCTGCCCGCTATGTCGCAATGTGTGGAGGCGCTCTGGAGG 368
Y 361 attataagacgagggcctatgagcccatgaggggctgacgtcagcaaacagcgc 420
D 369 ATTATAGACGAGGGCCCTATGGAGGCCCATGAGGGGGCTGACACGTCAAGCAACAGGC 428
Y 421 gcaggcgctgccacgcgcctttatttgcctgctacgaaaagttaaaagacattgagt 480
D 429 GCAGGGCTGCCACGCGCTTTATTTTGCCTGTCTACGAAAAGTTAAAAAAGACATTGAGT 488
Y 481 gatgtaataccactgggggcaatagcattatgccaatggt 522
D 489 GATGTAATCACCCTGGGGGCAATAGCCATATTGCCCAATGTT 530

RESULT 5
AX061229 AX061229 1322 bp DNA linear PAT 22-JAN-2001
LOCUS
DEFINITION Sequence 76 from Patent WO0078953.

AX061229
VERSION AX061229.1 GI:12406365
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1322)
AUTHORS Lal,P., Yang,J., Yue,H., Hillman,J.L., Tang,Y.T., Bandman,O.,
Burford,N., Baughn,M.R., Azimzal,Y., Lu,D.A., Au-Young,J. and
Patterson,C.
TITLE Human transport proteins
JOURNAL Patent: WO 0078953-A 76 28-DEC-2000;
Incyte Genomics, Inc. (US)
FEATURES
source 1..1322
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 328 a 376 c 351 g 267 t
ORIGIN

Query Match 65.6%; Score 454.4; DB 6; Length 1322;
Best Local Similarity 98.7%; Pred. No. 1.4e-67;
Matches 458; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Y 114 ccggggggcgcgcgagggcgagggcgctgagggcgcccggtgagcaagatccgga 173
D 2 CGGGGGCGCGCGGGGAGGCGGGGCTGCAGGCCCCCGGTAGCAAGATCCGGA 61
Y 174 ctccggggcgagtcagagggcgctgcccggctggagccactgtcaccacgcacatggtggc 233
D 62 CTCGGGCGCGACTACGAGGCGCTGCCGGCTGGAGCCACTGTCAACACGACATGGTGGC 121
Y 234 aggcgcgctgaggggactcctgagcactcgtgagtgtacccatcgactgctcaagac 293
D 122 AGCGCGCGTGGCAGGGATCCTGGAGCCTCGGTGATGATCCCATCGACTGCGTCAAGAC 181
Y 294 ccggatcagagctctacagcctgacccagctgcccgcctatcgcaatggttggaggccct 353
D 182 CCGGATGCAGGTCTACAGGCTGACCCAGCTGCCCGTATCGCAATGTGTGGAGGCCCT 241
Y 354 ctggaggattataagaacgagggcgcttatggagggccatgagggggcgtgaacgtcacgc 413
D 242 CTGAGGATTATAAGAACGAGGAGGCGCTATGGAGGCCCATGAGGGGGCTGAACGTCACAGC 301
Y 414 aacaggcgagggcgctgcccagccctttatttgcctgctagaaaaagattaaaaagac 473
D 302 AACAGGCGCAGGCGCTGCCACGCGCTTTATTTGCTGTCTACGAAAAGTTAAAAAGAC 361
Y 474 attgagtgatgtaatccaccctggggcaatagccatattgccaatggtgctggccgggtg 533
D 362 ATTGAGTGTATGTAATCCACCCTGGGGCAATAGCCATATTGCCAATGTTGCGCGCGGTG 421
Y 534 tgtggcaacattacttcatgctcagccatgaaacctgcggaag 577
D 422 TGTGGCAACATTACTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 465

RESULT 6
AL353719/c
LOCUS
DEFINITION Human DNA sequence from clone RP11-85A1 on chromosome 10, complete
sequence.
AL353719 AC007643
VERSION AL353719.10 GI:15787725
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 123160)
AUTHORS Ramsay,H.

REFERENCE AUTHORS

Rattus.
1 (bases 1 to 221062)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaraturunge,H.C., Are,J.R., Banks,T., Barbaria,J.,
Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowie,S., Brleva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havliak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
Hollins,B., Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Louisged,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M.,
Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Oragunye,N., Nickerson,E., Nwokenkwo,S.,
Ogulu,A., Okwuonu,G., Ogunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Pevens,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M.,
Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N.,
Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S.,
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 221062)
Worley,K.C.
Direct Submission
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Dec 20, 2001 this sequence version replaced gi:15627972.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information

Center project name: GE2D
Center clone name: CH230-24M6
----- Summary Statistics

Assembly program: Phrap; version 0.990329First call to
findPhrapList
Consensus quality: 200164 bases at least Q40
Consensus quality: 205303 bases at least Q30
Consensus quality: 209366 bases at least Q20
Estimated insert size: 209880; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-tp estimation
Quality coverage: 3.9x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 33 contigs. The true order of the pieces
* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
32851: contig of 32851 bp in length
32852: gap of unknown length
32951: contig of 16033 bp in length
48984: gap of unknown length
48985: contig of 14731 bp in length
63815: contig of 12090 bp in length
63915: gap of unknown length
76005: contig of 12090 bp in length
76105: gap of unknown length
88841: contig of 12736 bp in length
88842: gap of unknown length
100591: contig of 11650 bp in length
100691: gap of unknown length
100692: contig of 11011 bp in length
111703: gap of unknown length
111802: contig of 10862 bp in length
122665: gap of unknown length
122665: contig of 6947 bp in length
129711: contig of unknown length
129811: gap of unknown length
137247: contig of 7436 bp in length
137347: gap of unknown length
144676: contig of 7329 bp in length
144776: gap of unknown length
152124: contig of 7348 bp in length
152224: gap of unknown length
152225: contig of 6654 bp in length
158878: gap of unknown length
158979: contig of 4950 bp in length
163929: gap of unknown length
164029: contig of 5325 bp in length
169354: gap of unknown length
169453: contig of 6137 bp in length
175590: gap of unknown length
175591: contig of 4785 bp in length
180475: gap of unknown length
180476: contig of 4082 bp in length
180576: gap of unknown length
184657: contig of 4082 bp in length
184658: gap of unknown length
184758: contig of 5575 bp in length
190333: gap of unknown length
190433: contig of 3702 bp in length
194135: gap of unknown length
194235: contig of 2358 bp in length
196593: gap of unknown length
200530: contig of 3838 bp in length
200531: gap of unknown length
200631: contig of 2179 bp in length
202810: gap of unknown length
202910: contig of 2125 bp in length
205034: gap of unknown length
205134: contig of 2042 bp in length
207176: gap of unknown length
207277: contig of 1569 bp in length
208845: gap of unknown length
208946: contig of 2326 bp in length
211271: gap of unknown length
211371: contig of 1931 bp in length
213302: gap of unknown length
213403: contig of 2082 bp in length
215485: gap of unknown length
215485: contig of 1385 bp in length
215855: gap of unknown length
216370: contig of 1289 bp in length
217070: gap of unknown length
218359: contig of 1428 bp in length
218459: gap of unknown length
219887: contig of 1076 bp in length.
219987: Location/Qualifiers
1. .221062
/organism="Rattus norvegicus"

FEATURES
source

```

/db_xref="taxon:10116"
/chromosome="Rf1"
/clon="CH230-24M6"
BASE COUNT 58389 a 50577 c 49732 g 58877 t 3487 others
ORIGIN

Query Match 36.2%; Score 251; DB 2: Length 221062;
Best Local Similarity 91.4%; Pred. No. 2.le-33;
Matches 266; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 1 atgagttgagggcggggtgctggcggtgtggcgggggcgccgagggcccggg 60
|||||
Db 72865 ATGAGTTGAGGGCGGAGTGCGAGCGGCGTGGCGGGAGGACGAGTGTGGCCCGGG 72924
|||||

QY 61 cggagccccgggagtcggtgctgagcgggtggctgagcggggcggtggcgccgggg 120
|||||
Db 72925 CGGAGCCCCGGGAGTGGCGCTGCTGGACGGGTGGCTGCAGCGGGCGTGGCGCGGGG 72984
|||||

QY 121 gcggcgccggggagggcggggctgagcccccgcgggtacgacaagatccggactccggc 180
|||||
Db 72985 GCGCGCGCGGGAGCGGGGGCGCTATCAGACCCCTGTACGGCTGGATCCGGAGTCCGGT 73044
|||||

QY 181 ccggactacagggcgtgcggctggagccactgtaccaccgacatggtggcgagggcc 240
|||||
Db 73045 CCGGAATACGAAGCGCTGCCGGCTGGAGCCACTGTCTACCGACATGTTGGCGGCGCC 73104
|||||

QY 241 gtgacgggtacctggagcactgctgctgtaccctccactgactgctgcaag 291
|||||
Db 73105 GTGCGAGGATCTGGAGCATGTCGCTGATGTACCCGATCGACTCGGTCAAG 73155
|||||

RESULT 9
AC096351/c
LOCUS
DEFINITION Rattus norvegicus chromosome Rf1 clone CH230-24M6, WORKING DRAFT
SEQUENCE, 33 unordered pieces.
AC096351
VERSION AC096351.2 GI:17944054
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 221062)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbara,J.,
Benton,J., Bimge,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowle,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Dean,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
Hollins,B., Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,D.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Loulsegh,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhiney,E., McLeod,W.P., Meador,M.,
Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,

```

```

Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,
Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N.,
Sisson,I., Sodergren,E., Sonaik,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
Tang,H., Tansey,J., Taylor,C., Taylor,T., Teifrod,B., Thomas,N.,
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S.,
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 221062)
Worley,K.C.
Center: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
-----
Center project name: GE2D
Center clone name: CH230-24M6
----- Summary Statistics
-----
Assembly program: Phrap; version 0.990329First call to
findPhrapList
Consensus quality: 200164 bases at least Q40
Consensus quality: 205303 bases at least Q30
Consensus quality: 209366 bases at least Q20
Estimated insert size: 209880; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 3.9x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 33 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1
* 32851: contig of 32851 bp in length
* 32852 32951: gap of unknown length
* 32952 48984: contig of 16033 bp in length
* 48985 49084: gap of unknown length
* 49085 63815: contig of 14731 bp in length
* 63816 63915: gap of unknown length
* 63916 76005: contig of 12090 bp in length
* 76006 76105: gap of unknown length
* 76106 88841: contig of 12736 bp in length
* 88842 88941: gap of unknown length
* 88942 100591: contig of 11650 bp in length
* 100592 100691: gap of unknown length
* 100692 111702: contig of 11011 bp in length
* 111703 111802: gap of unknown length
* 111803 122664: contig of 10862 bp in length
* 122665 122764: gap of unknown length
* 122765 129711: contig of 6947 bp in length
* 129712 129811: gap of unknown length
* 129812 137247: contig of 7436 bp in length
* 137248 137348: gap of unknown length
* 137349 144676: contig of 7329 bp in length
* 144677 144776: gap of unknown length
* 144777 152124: contig of 7348 bp in length
* 152125 152224: gap of unknown length

```

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT


```
info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 15 Row: h Column: 21
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 7706149.
FEATURES
Location/Qualifiers
1..858
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:8764 IMAGE:3910567"
/tissue_type="uterus, leiomyosarcoma"
/clone_lib="NIH_MGC_71"
/lab_host="DH10B"
/notes="Vector: PCMV-SPORT6"
26..493
/codon_start=1
/product="Unknown (protein for MGC:8764)"
/protein_id="AAH15013.1"
/db_xref="GI:15929110"
/translation="MELRSGVSGQAVARRMDGSDRGGGKDATGSDYENLPTSA
VSTHGMVIMAGALEHSPVDSVKTRMQSLSPDKPAQYTSYIGALKIMRTGFWFR
PLRGVNMIMAGFAHAMFYFACIYENMKRTLNDVPHOGNHSLANGILKAFVWS"
BASE COUNT      215 a   220 c   231 g   192 t
ORIGIN

Query Match      24.8%; Score 172; DB 9; Length 858;
Best Local Similarity 69.1%; Pred. No. 8.5e-20;
Matches 235; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 183 ggactacgagcgctgcggctggagccactgtcaccagcaecatggtggcagggcgccgt 242
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 127 GGACTACGAGAACCTGCGGACATGAGCGCTCCGTGCTGCCACCCACATGACAGCAGGACCGAT 186
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 243 ggcaggatcctggagcactgcgtgatgtacccctgcactcgtcaagaccggatgca 302
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 187 GCCGGGATCTGGAGCACTCGGTGATGACCGGTGACTCGGTGACACACGACGATGCA 246
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 303 gagtacagcctgaccagctgcccgcgtatcgcgaatgtgttgaggccctctggagat 362
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 247 GAGTTTGAGTCCAGATCCCAAGCCAGTACACAAAGTATCTACGGAGCCCTCAAGAAAAAT 306
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 363 tataaagacgagcgccctgagagcccatgaggggctgaacgtcacagcaacaggcgc 422
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 307 CATGGCGGACCGAAGCGCTCTGGAGGCGCCTTGGGAGCGCTCAACGTCATGATCGGTGTC 366
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 423 agggcctgcgcacgccttattttgcctgcctacgaaagttaaaagacattgagtga 482
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 367 AGGGCGGCGCCATGCCATGCTATTTTGCTGCTGCTATGAARACATGAAGAGCCTTTAAATGA 426
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 483 tgtaatccacctgggggcaatagccatattgccaatggt 522
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 427 CGTTTTCACCAACCAAGAAACAGCCACCTAGCAGCGT 466
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 15
AC099368
LOCUS
DEFINITION
Rattus norvegicus clone CH230-84G11, *** SEQUENCING IN PROGRESS
***, 71 unordered pieces.
ACCESSION AC099368
VERSION AC099368.2 GI:17973175
KEYWORDS HTG; HTGS_PHASE1.

SOURCE
ORGANISM
Norway rat.
Rattus norvegicus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus
1 (bases 1 to 163205)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,
Albrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbara,J.,
Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burch,P., Burrell,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hoques,M., Holloway,C.,
Hollins,B., Homsif,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Loulsegh,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M.,
Mel,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,
Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,D., Peters,L., Pickens,R., Primus,E., Pu.L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,
Ruiz,S., Savery,G., Scherez,S., Scott,G., Shen,H., Shoohtari,N.,
Sisson,I., Sodergren,E., Sonaik,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K.,
Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S.,
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 163205)
Direct Submission
Worley,K.C.
Submitted (10-NOV-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Dec 20, 2001 this sequence version replaced gi:16901901.
-----
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
-----
Center project name: GHJ
Center clone name: CH230-84G11
-----
Summary Statistics
Assembly program: Phrap; version 0.990329First call to
findPhraplist
Consensus quality: 127058 bases at least Q40
Consensus quality: 138264 bases at least Q30
Consensus quality: 146548 bases at least Q20
Estimated insert size: 120308; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 1.9x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
```


THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2002, 08:05:43 ; Search time 664.31 Seconds
(without alignments)
1791.064 Million cell updates/sec

Title: US-09-870-113-5
Perfect score: 693
Sequence: 1 atggagttgagggcgggg.....ttacctgtttaccactag 693

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802.*
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES						
Result No.	Score	Query Match	Length DB ID	Description		
1	469.4	67.7	1336 21	AAC90457	Human uncoupling p	
2	458	66.1	1418 21	AAC74843	Human ORFX ORF398	
3	454.4	65.6	1322 22	AAF27733	Human transport pr	
c	440	63.5	1294 22	AAI60661	Human polynucleoti	
	401	57.9	2502 22	AAS03906	Human secreted pro	
6	398.6	57.5	1316 22	AAI58875	Human polynucleoti	
7	273	39.4	401 22	AAF66156	Novel human polynu	
8	233	33.6	27960 22	AAK69779	Human immune/haema	
9	233	33.6	27960 22	AAK73320	Human immune/haema	

10	200.4	28.9	1902	22	AAK521172	Human polynucleoti
11	196.6	28.4	1305	22	AAF59920	Human mitochondria
12	187.4	27.0	2562	22	AAH90057	Human bone marrow
13	181	26.1	339	22	AAK58672	Human immune/haema
14	174.2	25.1	1890	22	AAK94897	Human full-length
15	173.6	25.1	842	22	AAK92357	Human cDNA 5'-end
16	173.6	25.1	842	22	AAK93895	Human cDNA clone r
17	173.6	25.1	1814	22	AAK94871	Human full-length
18	172.6	24.9	1468	21	AAK94871	Human ORFX ORF2730
19	172	24.8	3028	22	AAC91320	Human polynucleoti
20	170.4	24.6	622	22	AAK91764	Human cDNA 5'-end
21	170.4	24.6	622	22	AAK93200	Human cDNA clone r
22	168.6	24.3	1913	22	AAH89944	Human bone marrow
23	160	23.1	1758	20	AAV84573	Human secreted pro
24	160	23.1	1758	22	ABA83356	Human secreted pro
25	132.8	19.2	1716	21	AAK77189	Human ORFX ORF3744
26	121	17.5	300	20	AAZ13026	Human gene express
27	121	17.5	710	20	AAZ15876	Human gene express
28	109.8	15.8	2837	22	AAK80493	Human immune/haema
29	104.8	15.1	483	24	ABI99871	Mouse ischaemic co
30	79.8	11.5	673	21	AAF12614	Aspergillus oryzae
31	72.4	10.4	2137	23	ABL08627	Drosophila melanog
32	67.8	9.8	3058	22	AAK83876	Human immune/haema
33	65.8	9.5	448	21	AAK77174	Human ORFX ORF2729
34	63.2	9.1	1000	21	AAA02484	Human colon cancer
35	59.2	8.5	349	22	AAI84504	Human polynucleoti
36	58.2	8.4	318	21	AAA38184	Primer used in the
37	58.2	8.4	320	21	AAA38186	Primer used in the
c	58.2	8.4	2188	20	AAZ77506	Human ovarian tumo
c	58.2	8.4	3198	20	AAZ02974	Human IL-1ra BAC c
40	58	8.4	2037	24	AAAG2787	cDNA sequence #574
41	57.6	8.3	320	21	AAA38185	Primer used in the
42	57.6	8.3	1698	23	AAK84408	DNA encoding novel
43	57.6	8.3	2049	23	AAK76210	DNA encoding novel
44	57.4	8.3	320	21	AAA38183	Primer used in the
45	57.2	8.3	1286	24	ABI99656	Mouse ischaemic co

ALIGNMENTS

RESULT	1
AAC90457	
ID	AAC90457 standard; cDNA; 1336 BP.
XX	
AC	AAC90457;
XX	
DT	12-MAR-2001 (first entry)
XX	
DE	Human uncoupling protein cDNA #6.
XX	
KW	Human; uncoupling protein; immunosuppressive; antiarthritic;
KW	antirheumatic; antiproliferative; cardiant; vasotropic;
KW	cerebroprotective; neuroprotective; antibacterial; ophthalmological;
KW	gastrointestinal; nephrotropic; gynaecological; vulnery; thrombolytic;
KW	gene therapy; cancer; wound; infectious disease; thrombosis; arthritis;
XX	infertility; ss.
XX	
OS	Homo sapiens.
XX	
FN	WO200061614-A2.
XX	
PD	19-OCT-2000.
XX	
PF	06-APR-2000; 2000WO-US09534.
XX	
PR	09-APR-1999; 99US-0128701.
PR	08-JUL-1999; 99US-0142821.
PR	18-AUG-1999; 99US-0149448.
PR	12-NOV-1999; 99US-0164751.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
XX	

Db	61	ccccccctacgagcgctgcgcggtgagccaactgtcaccacgcacatggtggcaggccgc	120
Qy	241	gtgcaggatccttgagcactcgtgatgatacccatgactgcgtcaagaccggatg	300
Db	121	gtgcaggatccttgagcactcgtgatgatacccatgactgcgtcaagaccggatg	180
Qy	301	cagagtctacagcctgaccagctgccgctatcgcaatgtgtggagccctctggagg	360
Db	181	cagagtctacagcctgaccagctgccgctatcgcaatgtgtggagccctctggagg	240
Qy	361	attataagacggaggccctatggagcccatgaggggcgtgaactcacagcaacaggc	420
Db	241	attataagacggaggccctatggagcccatgaggggcgtgaactcacagcaacaggc	300
Qy	421	gcaggcctgccacgccccttatttgctgctacgaaagttaaataagacattgagt	480
Db	301	gcaggcctgccacgccccttatttgctgctacgaaagttaaataagacattgagt	360
Qy	481	gatgatacaacctgggggcaatagccatgtgcgaatggtcggcccggtgtgtggca	540
Db	361	gatgatacaacctgggggcaatagccatgtgcgaatggtcggcccggtgtgtggca	420
Qy	541	acataacttcattgatgcagccatgaaccttcgggaagg	578
Db	421	acataacttcattgatgcagccatgaaccttcgggaagg	458

RESULT 3

AAE27733
ID AAE27733 standard; cDNA: 1322 BP.

AC AAF27733:

DT 28-MAR-2001 (first entry)

DE Human transport protein TPPT-33 coding sequence.

Human; transport protein; TPPT; transport disorder; metabolic disorder;
KW
KW
neurological disorder; cardiovascular disorder; reproductive disorder;
immune disorder; cancer; ss.
KW

OS Homo sapiens.

PN WO200078953-A2.

28-DEC-2000.

16-JUN-2000; 2000WO-US16668.

AA 17-JUN-1999; 99US-0139923.

PR 10-AUG-1999; 99US-0148177.

PR 18-AUG-1999; 99US-0149357.

PR 28-OCT-1999; 99US-0162287.

XX
PA (TNCV-) TNCVTE GENOMICS INC

XX
PI Lal P., Yang J., Yue H., Hillman JL, Tang YT, Bandman O, Burford N;
PI Baughn MR, Azimzai Y, Lu DAM, Au-Young J, Patterson C;
XX

WPI: 2001-041424/05.

DR P-PSDB; AAB60113.

Isolated polypeptide with a human transport protein sequence is useful for the diagnosis, prevention and treatment of disorders associated with the immune, reproductive and cardiovascular systems

PS Claim 5; Page 160; 165pp; English.

CC The present invention provides the protein and coding sequences for 43
CC novel human transport proteins (designated TPPs). These can be used in
CC the diagnosis and treatment of transport, metabolic, neurological,
CC reproductive, cardiovascular and immune disorders, and cell proliferative

CC disorders such as cancer.

Sequence 1322 BP; 328 A; 376 C; 351 G; 267 T; 0 other;
AA

Query Match 65.6%; Score 454.4; DB 22; Length 1322;
Best Local Similarity 98.7%; Pred. No. 5.4e-88;
Matches 458: Conservative 0; Mismatches 6; Indels 0;

Qy	114	ccgggggcccggcgccgggagcgccgggctcagggcccccgtacgacaaagatccgga	173
Db	2	cgggggccggcgcgccgggagcgccgggctcagggcccccgtacgacaaagatccgga	61
Qy	174	ctcgggcccgagctacagagcgctgcgctgagcactgtccacacacatgtgac	233
Db	62	ctcgggcccgagctacagagcgctgcgctgagcactgtccacacacatgtgac	121
Qy	234	aggcgccgtggcaggatcctgagcaactcgtgatgccccatcgactcgctcaagac	293
Db	122	aggcgccgtggcaggatcctgagcaactcgtgatgccccatcgactcgctcaagac	181
Qy	294	ccgagtcagagtcacagctgaaccagctgccgctatcgaaatgtgttgaggccct	353
Db	182	ccgagtcagagtcacagctgaaccagctgccgctatcgaaatgtgttgaggccct	241
Qy	354	ctggaggtatataagaacgaggccctatggagcccatgaggggctgaactcagac	413
Db	242	ctggaggtatataagaacgaggccctatggagcccatgaggggctgaactcagac	301
Qy	414	aacaggcgagggcctgccacgccttatttgcctgtctcgaaaaagttaaaaaagac	473
Db	302	aacaggcgagggcctgccacgccttatttgcctgtctcgaaaaagttaaaaaagac	361
Qy	474	attgagtgtatcaatcacccctggggcaatagccatttgcgaatgtgtgcgccgggtg	533
Db	362	attgagtgtatcaatcacccctggggcaatagccatttgcgaatgtgtgcgccgggtg	421
Qy	534	tgtggcaacttaactcatgatgcagccatgaacctgcggaag	577
Db	422	tgtggcaacttaactcatgatgcagccatgaacctgcggaag	465

RESULT

AAI60661/C

ID AAI60661 standard; cDNA; 1294 BP.

AC AAI60661;

DT 22-OCT-2001 (first entry)

Human polynucleotide SEO ID NO 4650.

Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.

OS Homo sapiens.

AA PN WO200153312-A1.

XX
PD
26-JUL-2001.XX
PF 26-DEC-2000: 2000WO-IIS34263XX
PR 21-JAN-2000: 2000US-0488725

PR 23-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598043
PR

PK 05-JUL-2000; 2000US-0398042.
PR 19-JUL-2000; 2000US-0620312.

PR 03-AUG-2000: 2000US-0653450.

PR 14-SEP-2000; 2000US-0662191..

Db 66 cggagactcggccgactacagagcgctgceggctggagcgaactgtccaccgacatg 125
QY 229 gtgcaggcgccgtggcaggatcctggagcactgcgtgatgtacccctcgactcgctc 288
Db 126 gtgcaggcgccgtggcaggatcctggagcactgcgtgatgtacccctcgactcgctc 185
QY 289 aagaccggatgcagagtgatcagcctgacagcctgaccagctgcccgcgtatcgcaatgtgtggag 348
Db 186 aagaccggatgcagagtgatcagcctgacagcctgaccagctgcccgcgtatcgcaatgtgtggag 245
QY 349 gcccttgaggattataaagcaggagggcctatgaggcccatgagggggctgaacgtc 408
Db 246 gcccttgaggattataaagcaggagggcctatgaggcccatgagggggctgaacgtc 305
QY 409 aagcaaacaggcgagcgctgcccacgcccctttatttgcctgctacgaaagttaaaa 468
Db 306 aagcaaacaggcgagcgctgcccacgcccctttatttgcctgctacgaaagttaaaa 365
QY 469 aagacattgagtgatgaatccaccctggggcaatagccatattgccaatggt 522
Db 366 aagacattgagtgatgaatccaccctggggcaatagccatattgccaatggt 419
RESULT 6
AAI58875
ID AAI58875 standard; cDNA; 1316 BP.
XX AC AAI58875;
XX DT 22-OCT-2001 (first entry)
XX Human polynucleotide SEQ ID NO 1078.
DE Human; neurotropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX OS Homo sapiens.
XX PN WO200153312-A1.
XX PD 26-JUL-2001.
XX PF 26-DEC-2000; 2000WO-US34263.
XX PR 21-JAN-2000; 2000US-0488725.
XX PR 25-APR-2000; 2000US-0552317.
XX PR 09-JUL-2000; 2000US-0598042.
XX PR 19-JUL-2000; 2000US-0620312.
XX PR 03-AUG-2000; 2000US-0653450.
XX PR 14-SEP-2000; 2000US-0662191.
XX PR 19-OCT-2000; 2000US-0693036.
XX PR 29-NOV-2000; 2000US-0727344.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX WPI: 2001-442253/47.
DR P-PSDB; AAM39719.
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX Claim 1; SEQ ID NO 1078; 10078pp; English.
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and

CC the encoded polypeptides (AAM38642-AAM42213) with neurotropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemia and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX SQ Sequence 1316 BP; 335 A; 364 C; 337 G; 280 T; 0 other;
Query Match 57.5%; Score 398.6; DB 22; Length 1316;
Best Local Similarity 95.5%; Pred. No. 4.4e-76;
Matches 421; Conservative 0; Mismatches 19; Indels 1; Gaps 1;
QY 138 cggggcctcaggcccccggtacgacaaatccggactccgcccggactacagcgct 197
Db 15 cgaggcctgcaggcccccggtacgacaaatccggactccgcccggactacagcgct 74
QY 198 gccggctgagccactgtcaccacgcacatg-gtggcaggcgccgtggcaggactcctgg 256
Db 75 gtccgctgagccactgtcaccacgcacatgtgggtatgctgcgcgagccatggtcgtg 134
QY 257 agcaactgcgtgatgtaccccatcgactgcgtcaagaccgcggatgcagagctcacagcctg 316
Db 135 agcaactgcgtgatgtaccccatcgactgcgtcaagaccgcggatgcagagctcacagcctg 194
QY 317 accagctgccgctatcgcaatgtgtggaggccctctggaggattataagaaacggagg 376
Db 195 accagctggcgctatcgcaatgtgtggaggccctctggaggattataagaaacggagg 254
QY 377 gcctatggagcccatgagggggctgaacgtcacacaacagcagcgccctgccacg 436
Db 255 gcctatggagcccatgagggggctgaacgtcacacaacagcagcgccctgccacg 314
QY 437 ccctttatttgcctgctacgaaaagttaaaaaagacattgagtgaatccaccctg 496
Db 315 ccctttatttgcctgctacgaaaagttaaaaaagacattgagtgaatccaccctg 374
QY 497 ggggcaatagccatattgccaatgtgcggccgggtgtgtggcaacattacttcattg 556
Db 375 ggggcaatagccatattgccaatgtgcggccgggtgtgtggcaacattacttcattg 434
QY 557 cagccatgaacctgcggaag 577
Db 435 cagccatgaacctgcggaag 455
RESULT 7
AAF66156
ID AAF66156 standard; cDNA; 401 BP.
XX AC AAF66156;
XX DT 09-APR-2001 (first entry)
XX Novel human polynucleotide, SEQ ID NO: 1912.
DE Human; cytostatic; gene therapy; colon cancer; prostate cancer;
KW breast cancer; lung cancer; cancer detection; ss.
XX OS Homo sapiens.
XX PN WO200102568-A2.
XX

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer; cytosolic; gene therapy; vaccine; metastasis; ds.
KW OS Homo sapiens.
XX WO200157182-A2.
XX 09-AUG-2001.
XX 17-JAN-2001; 2001WO-US01354.
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233053.
PR 14-SEP-2000; 2000US-0233064.
PR 21-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 26-SEP-2000; 2000US-0234998.
PR 27-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 29-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-024617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249219.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.

polypeptide encoded by it are useful in the treatment of various immune deficiencies and disorders. The deficiencies and disorders may be genetic, may be caused by a viral (e.g. HIV), bacterial or fungal infection, or may result from an autoimmune disorder, a coagulation disorder (e.g. haemophilia), inhibition of tumour cell proliferation, suppression of an inflammatory response or treatment of a nervous system disorder such as Alzheimer's disease. Detection of the presence or increased expression of the polynucleotide or the protein it encodes is useful for the diagnosis and/or prognosis of one or more types of cancer. The polynucleotide and polypeptide can be used as nutritional sources or supplements and in the screening of chemical compounds as potential drugs.

Sequence 2562 BP: 623 A; 660 C; 627 G; 652 T; 0 other;

Query Match	27.0%	Score 187.4;	DB 22;	Length 2562;
Best Local Similarity	68.2%;	Pred. No. 5.6e-31;		
Matches 260;	Conservative	0;	Mismatches 121;	Indels 0;
Gaps	0;			

[illegible]

RESULT	13	
AAK58672		
ID	AAK58672	standard; cDNA; 339 BP.
XX		
AC	AAK58672;	
XX		
DT	06-NOV-2001	(first entry)
XX		
DE	Human immune/haematopoietic antigen encoding cDNA	SEQ ID NO:3732.
XX		
KW	Human; Immune; haematopoietic; immune/haematopoietic antigen; cancer;	
KW	cystostatic; gene therapy; vaccine; metastasis; ss.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200157182-A2.	
XX		
PD	09-AUG-2001.	
XX		
PF	17-JAN-2001; 2001WO-US01354.	
XX		
PR	31-JAN-2000; 2000US-0179065.	
PR	04-FEB-2000; 2000US-0180628.	
PR	24-FEB-2000; 2000US-0184664.	

PR	02-MAR-2000;	2000US-0186350.
PR	16-MAR-2000;	2000US-0198974.
PR	17-MAR-2000;	2000US-0190076.
PR	18-MAR-2000;	2000US-0198123.
PR	19-MAY-2000;	2000US-0205515.
PR	19-JUN-2000;	2000US-0209467.
PR	28-JUN-2000;	2000US-0214886.
PR	30-JUN-2000;	2000US-0215135.
PR	07-JUL-2000;	2000US-0216647.
PR	07-JUL-2000;	2000US-0216880.
PR	11-JUL-2000;	2000US-0217487.
PR	11-JUL-2000;	2000US-0217496.
PR	14-JUL-2000;	2000US-0218290.
PR	26-JUL-2000;	2000US-0220963.
PR	26-JUL-2000;	2000US-0220964.
PR	14-AUG-2000;	2000US-0224518.
PR	14-AUG-2000;	2000US-0224519.
PR	14-AUG-2000;	2000US-0225213.
PR	14-AUG-2000;	2000US-0225214.
PR	14-AUG-2000;	2000US-0225266.
PR	14-AUG-2000;	2000US-0225267.
PR	14-AUG-2000;	2000US-0225268.
PR	14-AUG-2000;	2000US-0225270.
PR	14-AUG-2000;	2000US-0225247.
PR	14-AUG-2000;	2000US-0225757.
PR	14-AUG-2000;	2000US-0225758.
PR	14-AUG-2000;	2000US-0225759.
PR	18-AUG-2000;	2000US-0226279.
PR	22-AUG-2000;	2000US-0226681.
PR	22-AUG-2000;	2000US-0226688.
PR	23-AUG-2000;	2000US-0227182.
PR	23-AUG-2000;	2000US-0227709.
PR	30-AUG-2000;	2000US-0228924.
PR	01-SEP-2000;	2000US-0229287.
PR	01-SEP-2000;	2000US-0229343.
PR	01-SEP-2000;	2000US-0229344.
PR	01-SEP-2000;	2000US-0229345.
PR	05-SEP-2000;	2000US-0229509.
PR	05-SEP-2000;	2000US-0229513.
PR	06-SEP-2000;	2000US-0230437.
PR	06-SEP-2000;	2000US-0230438.
PR	08-SEP-2000;	2000US-0231242.
PR	08-SEP-2000;	2000US-0231243.
PR	08-SEP-2000;	2000US-0231244.
PR	08-SEP-2000;	2000US-0231413.
PR	08-SEP-2000;	2000US-0231414.
PR	08-SEP-2000;	2000US-0232080.
PR	08-SEP-2000;	2000US-0232081.
PR	12-SEP-2000;	2000US-0232168.
PR	14-SEP-2000;	2000US-0232363.
PR	14-SEP-2000;	2000US-0232364.
PR	14-SEP-2000;	2000US-0232398.
PR	14-SEP-2000;	2000US-0232399.
PR	14-SEP-2000;	2000US-0232400.
PR	14-SEP-2000;	2000US-0232401.
PR	14-SEP-2000;	2000US-0232403.
PR	21-SEP-2000;	2000US-02323065.
PR	21-SEP-2000;	2000US-0232423.
PR	21-SEP-2000;	2000US-0234274.
PR	25-SEP-2000;	2000US-0234999.
PR	25-SEP-2000;	2000US-0234998.
PR	26-SEP-2000;	2000US-0235484.
PR	27-SEP-2000;	2000US-0235834.
PR	27-SEP-2000;	2000US-0235835.
PR	29-SEP-2000;	2000US-0236327.
PR	29-SEP-2000;	2000US-0236367.
PR	29-SEP-2000;	2000US-0236368.
PR	29-SEP-2000;	2000US-0236369.
PR	29-SEP-2000;	2000US-0236370.
PR	02-OCT-2000;	2000US-0236802.
PR	02-OCT-2000;	2000US-0237037.
PR	02-OCT-2000;	2000US-0237038.
PR	02-OCT-2000;	2000US-0237039.

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2002, 06:22:39 ; Search time 4881.77 Seconds
(without alignments)
1915.983 Million cell updates/sec

Title: US-09-870-113-5
Perfect score: 693
Sequence: 1 atgagtgtagggcgggg.....ttacctgtttaccactag 693

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	532.6	76.9	888	9	AL530804
2	450.2	65.0	817	10	BI103329
3	438.4	63.3	1131	10	BG295496
4	419.4	60.5	573	10	BI541703
5	409.8	59.1	697	10	BI854643
6	398.2	57.5	967	10	BI739710
7	386.6	55.8	645	10	BG964218
8	367.2	53.0	397	9	BE012485
9	354.6	51.2	574	10	BI185580
10	346.4	50.0	681	9	AV704087
11	344.6	49.7	598	10	BE913718
12	323	46.6	560	10	BI794616
13	320.2	46.2	480	10	BI045863
14	302.2	43.6	336	9	AW326482
15	298.4	43.1	546	10	BM488747
16	292	42.1	546	9	BE014800
17	290.8	42.0	900	10	BG325430

18	288.6	41.6	747	10	BI827322
19	282.8	40.8	501	9	AW211366
20	274.4	39.6	916	10	BI757180
21	255.8	36.9	622	10	BJ039504
22	244.6	35.3	722	10	BI876051
23	243.6	35.2	417	9	AA104365
24	240.4	34.7	855	10	BI412175
25	240.2	34.7	261	10	BI338246
26	231.4	33.4	701	10	BI831943
27	222.6	32.1	421	10	BI448355
28	207.4	29.9	678	9	BB654057
29	203.2	29.3	561	9	AI942584
30	196.6	28.4	684	9	AI133696
31	196.6	28.4	750	10	BI831750
32	194.6	28.1	729	10	BI460337
33	194	28.0	567	10	BJ035232
34	193.6	27.9	512	9	AW210205
35	191.6	27.6	453	10	BF916224
36	190.2	27.4	612	10	BJ012200
37	187.4	27.0	1295	11	AK015770
38	186.8	27.0	803	10	BG777404
39	184.6	26.6	217	9	AW346915
40	182.8	26.4	331	10	BI449976
41	181.6	26.2	1209	11	AK006155
42	177.8	25.7	666	9	BB644566
43	177.2	25.6	513	10	BI466991
44	174.2	25.1	866	10	BF797743
45	172	24.8	592	10	BM313240

ALIGNMENTS

RESULT 1

AL530804

LOCUS

DEFINITION

AL530804

ACCESSION

AL530804

VERSION

AL530804.1

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 888)

AUTHORS

Li.W.B., Gruber.C., Jessee.J. and Polayes.D.

TITLE

Full-length cDNA libraries and normalization

JOURNAL

Unpublished (2001)

COMMENT

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

Location/Qualifiers

1..888

source

organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="CSDD008YH17"

/sex="male"

/tissue_type="neuroblastoma cells"

/lab_host="DH10B"

/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA

was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and

cloned into the Not I and Eco RV sites of the pCMVSPORT 6

vector. Library was normalized. Library was constructed

by Life Technologies. Contact : Feng Liang Life

Technologies, a division of Invitrogen 9800 Medical Center

Drive Rockville, Maryland 20850, USA Fax : (1) 301 610

8371 Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com"

168 a 244 c 322 g 139 t 15 others

BASE COUNT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. NIH-MGC <http://mgc.nci.nih.gov/>. National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999). Contact: Robert Strausberg, Ph.D. Email: cgapbs-re@mail.nih.gov Tissue Procurement: The Cepko Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov> Plate: LLAM10377 row: c column: 07 High quality sequence stop: 753.

REFERENCE 1 (bases 1 to 1131)
AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES
source

1..1131
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:4504638"
/clone_lib="NIH_MGC_94"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 3.3 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."
261 a 335 c 375 g 160 t

BASE COUNT
ORIGIN

Query Match 63.3%; Score 438.4; DB 10; Length 1131;
Best Local Similarity 91.9%; Pred. No. 2e-75; Indels 3; Gaps 3;
Matches 496; Conservative 0; Mismatches 41;

QY 41 ggccggcgagggccggcgagcccgggagctcgccgctgctggcgaggtg-gctg 99
DB 1 GACCAGCTGCTGGGCGCGGGCGAGCCCCGGGAGTCGGCGCTGTGGACGGGTGCTG 60
QY 100 cagcg-gggcgtggccggggggcgggcgggcgagggcgggcgctgagcccccgt 158
DB 61 CAGCGTGGCGTGGCGGGGGCGCGGTGGCGGGAGCGGGCGCTATCAGCCCCCTGT 120
QY 159 acgacaagatccgactccggccggactacagcg-gctgcggctgagccactgtca 217
DB 121 ACGGTGGATCCGGAGTCCGGCCCGGAATACGAAGCAGCGTCCGGCTGGAGCCACTGTCA 180
QY 218 ccacgcacatggtggcagcgccgtggcaggatccctggagcactgctgagtacccca 277
DB 181 CCACGCACATGTGGGGCGCGCGTGGCAGGATCCTGGAGATTGCGTGATGTACCCGA 240
QY 278 tgcactgcgaagaccggatgcagagtctcagcgtgaccagctgcccgcctatcgca 337
DB 241 TCAGTGTCTAAGACCCGGATGCGAGCGCTACAGCCTACCCAGCGCCGCTATCGGA 300
QY 338 atggtgtgagggccctctgagagattataagaacgaggggacctatgagggccatgagg 397
DB 301 ACGTGTGGAGGCTCTCTGGGAATCATGAGGACAGAGGGCCCTGTGGAGGCCCATGCGGG 360
QY 398 ggctgaactcacgacaacagcgagggccctgcccacgccccttatttggctgtaag 457
DB 361 GGCTGAACGTACAGCAACAGCGGGGGCCCTGCCACGCCCTCTATTGCTGTGCTAG 420
QY 458 aaaagttaaaagacattgagtgtgtaataccaccctgggggcaatgccaattattgcca 517
DB 421 AAAAGTAAAAAAGACATTGAGTGACGTAATCACCCAGGGGGCAATAGCCATATTGCCA 480
QY 518 atgggtgcccgggtgtgtggcaacattacttctatgtagcagccatgaacctgcggaag 577
DB 481 ATGGTGCACCCGGATGTGTGGCGAGATTACTTATGATGACGCCATGAATCCAGCGGAAG 540

RESULT 4
BI541703

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

TITLE
JOURNAL
MEDLINE
COMMENT

FEATURES
source

1..573
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="MARC 1BOV"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI; Library made from pooled tissue from lymph node, ovary, fat, hypothalamus, and pituitary."
121 a 166 c 176 g 110 t

BASE COUNT
ORIGIN

Query Match 60.5%; Score 419.4; DB 10; Length 573;
Best Local Similarity 95.4%; Pred. No. 9.4e-72; Indels 0; Gaps 0;
Matches 432; Conservative 0; Mismatches 21;

QY 125 gggcggggagggcgggcgctgagcccccgggtacgacaagatccggactccggccgg 184
DB 1 GCGGGGGGAGCGCGGGCGCTCAGGCCCTGTAGCAGGATCCGGACTCCAGGCCGG 60
QY 185 actagagggcgtgcccggctggagccactgtcacacgcacatggtggcaggcgcgtgg 244
DB 61 ACTACAGCGCGTGGCGGTGGAGTACTGTCCACGCACATGTTGGGGGGCGCGCTGG 120
QY 245 caggatccctgagcactgctgtagtaccctatcgactgctcaagaccggatgcaga 304
DB 121 CAGGGATCTTGAGGCACTGCGGTGATGTACCCGCTGACTGCGTCAAGCCCGGATGCAG 180
QY 305 gtctacagctgacccagctgccgctatcgcaatgtgttgaggccctcttgaggattca 364
DB 181 GCCTACAGCGGATCCAGCGCCCGCTATCGCAATGTGTGGAGGCCCTCTTGAGGATTA 240
QY 365 taagaacggaggccctatggaggcccatgaggggctgaactcagcaacagcgcgag 424

BI541703 573 bp mRNA linear EST 30-AUG-2001
455441 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.

BI541703
BI541703.1 GI:15382815
EST.
SOURCE
Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
1 (bases 1 to 573)

Smith, T. P. L., Grosse, W. M., Freking, B. A., Roberts, A. J., Stone, R. T., Casas, E., Wray, J. E., White, J., Cho, J., Fahrrenkrug, S. C., Bennett, G. L., Heaton, M. P., Laegreid, W. W., Rohrer, G. A., Chitko-Mckown, C. G., Pertea, G., Holt, I., Karamycheva, S., Liang, F., Quackenbush, J. and Keeler, J. W.

Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)

21180013
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390

Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -mismatch 12 options.

PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCTCCAGTCAGCAGC
Plate: 121 row: O column: 24
Seq primer: APTTAGTGCACACTATAG.

Location/Qualifiers
1..573

BASE COUNT 154 a 278 c 348 q 187 t

Note: this is a NIH_MGC Library.

Query Match	57.5%	Score 398.2	DB 10	Length 967
Best Local Similarity	89.8%	Pred. No. 1.3e-67		
Matches 473	Conservative 0	Mismatches 48	Indels 6	Gaps 4
QY 1	atcgagttgagggcgggggtgctggcggtgtgacggggggggcggcgcgcagggcccggg 60			
Db				
111	ATGAGATTGGAGGGCGGAGTGCAGGCGCGTGGCGGAGGACACAGCTGCTGGGCCCGGG 170			
QY 61	cggagccccggggagtcggcgctgtggaacgggtggctgc--agcgggggcgtgggcccgg 117			
Db				
171	CGGAGCCCCGGGGAGTCGGCGCTGCTGACAGGGTGGCTGCTAGCGGTGGCGTGGGCTGG 230			
QY 118	ggggccggcgggggggggcgggcgccgtgcagggcccccggtac-gacaagatccggactc 176			
Db				
231	GGGCGCGGCGGGGAGCGGGGCCCTATTACAGCCCCCTGTACTGCTGGATCCGGAGTC 290			
QY 177	cggcccgagctacagcgcgtgcggctgagcgcactgtcaccacgcacatgggtggcagg 236			
Db				
291	CGGCCCCGAATACGAGCGCTGCCGCTGGAGCCACTGTCCACCACGCACAT-GTGGCGGG 349			
QY 237	cgcgctggcagggatcctggagcactgcgtgatgtacccccactgcactcgtccaagaccg 296			
Db				
350	CGCGTGGCAGGGATCCTGGAGCATTCGGTGATGTACCCGATCGACTGCGTCAAGACCCG 409			
QY 297	gatcgaggtctacagcctgacccagctgcccggtatcgcaatgtgtggaggccctctg 356			
Db				
410	GATGCAGAGCCTACAGCCTGACCCAGCGCCCGCTATCGGAACGCTGTGGAGGCTCTCTG 469			
QY 357	gagattataagaacgagggcctatgagcccatgaggggctgaactcagagcaac 416			
Db				
470	GAGAAATCATGAGGACAGGGCCCTGTGGAGGCCCATCGGGGGGTGAACGTCACAGCAAC 529			
QY 417	aggcgagggcctgcccaacgccttattttgctgctacgaaaagttaaaaaagacatt 476			
Db				
530	AGGCGCGGGCCTGCCACGCCCTCTATTTTGCTCTACGAAAAAGTTAAAAAGACATT 589			
QY 477	gagtgtgtatccaccctgggggc-aatagccaatattgccaatgggt 522			
Db				
590	GAGTGACGTAAATCACCCAGGGGCAATAGCCATATTGCCAATGGT 636			

RESULT	7
BG964218	
LOCUS	linear EST 12-JUN-2001
DEFINITION	602828996F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4903815 '5'; mRNA
ACCESSION	BG964218
VERSION	BG964218
KEYWORDS	GI:14351855
SOURCE	EST.
ORGANISM	Mus musculus house mouse. Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	1 (bases 1 to 645)
TITLE	NIH-MGC http://mgc.nci.nih.gov/ .
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgaqbs-r@mail.nih.gov Tissue Procurement: Jeffrey E. Green, M.D. CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM10988 row: p column: 24 High quality sequence start: 2

High quality sequence stop: 632.	
Location/Qualifiers	
1. .645	
/organism="Mus musculus"	
/strain="FVB/N"	
/db_xref="taxon:10090"	
/clone_image="IMAGE:4983815"	
/clone_lib="NCI-CGAP_Co24"	
/lab_host="DH10B (T1 phage-resistant)"	
/note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.6 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP library."	
147 a 189 c 193 g 116 t	
BASE COUNT	
ORIGIN	

Query Match	55.8%	Score 386.6	DB 10	Length 645
Best Local Similarity	91.3%	Pred. No. 2.2e-65		
Matches 410	Conservative 0	Mismatches 39	Indels 0	Gaps 0
Qy 129	cggggagggccggcctgcagcccccgggtacgacaagatcgcgactcgcgcccggacta	188		
Db 4	CGCGTAGCGGGGCCATCATGACCCCTGTACGGCTGGATCCGGACTCGGCCCGGAATA	53		
Qy 189	cgaggcgtgcggcgtggagccactgtccaccacgcacatggtggcaggcgcgtgtgcagg	248		
Db 64	CGAAGCGGTGCGCGCTGGAGCCACTGTCCACCACGCACATGGTGGCGGGCGCGTGGCAGG	123		
Qy 249	gattcctgagcactgcgtgatgtaccctcatcgactgcgtcaagaccggatgcagagtct	308		
Db 124	GATCCTGGAGCATTCGCTGATGTACCCGATCGACTGCGTCAAGACCCCGGATGCAGAGCT	183		
Qy 309	acagcctgaccagctgcccgctatcgcaatgtgttgaggccctctggaggattataag	368		
Db 184	ACAGCCTGACCACGCCGCCGCTATCGGAACGCTGTTGGAGGCTCTCTGGAGAAATCATGAG	243		
Qy 369	aacggagggcctatggagggccatgagggcgtgaacgtcacagcaacagcgcagggcc	428		
Db 244	GACAGAGGGCTGTGGAGGCCCATGCGGGGCTGAACGTACAGCCACAGGCGCGGGGCC	303		
Qy 429	tgccacgcgccttattttgctgctactcgaagaaagttaaaaaagacattgagtgtaat	488		
Db 304	TGCCACGCGCTCATTTTGTGCTGCTACGAAAGTTAAAAAGACATTTGAGTGAGCTAAT	363		
Qy 489	ccacctggggcgaatagccatattgccaatgtgtgcggcgggtgtgtgcaacattact	548		
Db 364	CCACCCAGGGGCCAATAGCCATATTGCCAATGGTGCAGCCGGATGTGTGGCGACATTTACT	423		
Qy 549	tcattatgcagccatgaacctgcggaag	577		
Db 424	TCATGATGCAGCCATGAATCCAGCGGAAG	452		
RESULT 8				
BE012485				
LOCUS	BE012485	397 bp	mrna	linear
DEFINITION	121949 MARC lPIG Sus scrofa cDNA 5', mRNA sequence.			EST 09-JUL-2000
ACCESSION	BE012485			
VERSION	BE012485.1	GI:8273406		
KEYWORDS	EST.			
SOURCE	plg.			
ORGANISM	Sus scrofa			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
REFERENCE	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.			
AUTHORS	1 (bases 1 to 397)			
	Fahrenkrug, S.C., Freking, B.A., Rohrer, G.A., Smith, T.P.L., Casas, E.,			
	Stone, R.T., Heaton, M.P., Grosse, W.M., Bennett, G.A., Laegreid, W.W.			
	and Keeler, J.W.			
TITLE	Design and use of two pooled tissue normalized cDNA libraries for			
	EST discovery in swine			
JOURNAL	Unpublished (2000)			
COMMENT	Contact: Smith TPL			


```
RESULT 10
AV704087
LOCUS      681 bp      mRNA      linear      EST 09-OCT-2000
DEFINITION AV704087 ADB Homo sapiens cDNA clone ADBAF05 5', mRNA sequence.
ACCESSION  AV704087
VERSION     AV704087.1 GI:10721407
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
REFERENCE  1 (bases 1 to 681)
AUTHORS   Peng, Y., Song, H., Huang, Q., Huang, C., Gu, Y., Yang, Y., Gao, G., Xiao
           H., Xu, X., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z.,
           Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu
           G., Hu, R., Chen, J., Chen, Z. and Han, Z.
TITLE     Homo sapiens cDNA ADB clones
JOURNAL   Unpublished (2000)
COMMENT   Contact: Zeguang Han
           Chinese National Human Genome Center at Shanghai
           351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
           201203, P. R. China
           Tel: 86-21-50801919(ex.45)
           Fax: 86-21-50801922
           Email: hanzg@chgc.sh.cn
           This clone is available at CHGC in Shanghai.
FEATURES   source
           1..681
           /organism="Homo sapiens"
           /db_xref="taxon:9606"
           /clone="ADBAFE05"
           /clone_lib="ADB"
           /tissue_type="Adrenal gland"
           /dev_stage="Adult"
           /lab_host="SOUR"
           /note="vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
           XhoI"
BASE COUNT 149 a 209 c 205 g 117 t 1 others
ORIGIN
Query Match 50.0%; Score 346.4; DB 9; Length 681;
Best Local Similarity 99.7%; Pred. No. 1.4e-57;
Matches 347; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 116 gggggccgcgcggggagccggggcctgcagcccccgggtacgacaagatccgaact 175
Db 1 GGGGGCCGCGGGGAGCCGGGGCCCTCAGCCCCCGGTACGCAAGATCCGGACT 60
Qy 176 ccggcccgactacagcgctgcggctggagccactgtcaccacgcacatggtggcag 235
Db 61 CCGGCCCGACTACGAGCGCTGCGGGCTGGAGCCACTGTCCACCGCACATGTGTGGCAG 120
Qy 236 gcgcctggcaggatcctggagcactgcgtgtatgtatcccccatgcactgcgtcaagacc 295
Db 121 GCGCGAGCGAGGATCCTGGGACACTCGGTGATGTACCCCATCGACTGCGTCAAGACCC 180
Qy 296 ggatgcagagctcagcctgaccagctgcccagctgcccgtatcgcaatgtgtggagccctct 355
Db 181 GGATGCAGAGTCTACAGCTCACCACAGCTGCCCGCTATFCGCAATGTGTGGAGGCCCTCT 240
Qy 356 ggaggattaaagcaggagccctatggagcccatggagggcctgaacgtcacagcaa 415
Db 241 GGAGGATTTATGACGAGGAGGCGCTATGAGAGCCCATGAGGGGCTGAACGTCAAGCAA 300
Qy 416 caggcgaggccctgccacacccctttattttgctgctacgaaaaagt 463
Db 301 CAGGCGCAGGGCCTGCCACACGCCCTTTATTTTGGCTGCTACGAAAAGT 348
RESULT 11
BE913718
LOCUS      598 bp      mRNA      linear      EST 29-SEP-2000
DEFINITION BE913718 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:3969256 5',
           mRNA sequence.
ACCESSION  BE913718
VERSION     BE913718.1 GI:10411618
KEYWORDS   EST.
SOURCE     house mouse.
ORGANISM   Mus musculus
REFERENCE  1 (bases 1 to 598)
AUTHORS   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE     NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL   National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT   Unpublished (1999)
           Contact: Robert Strausberg, Ph.D.
           Email: cgapbs-r@mail.nih.gov
           Tissue Procurement: Gilbert Smith, Ph.D.
           cDNA Library Preparation: Life Technologies, Inc.
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Incyte Genomics, Inc.
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLAM9146 row: o column: 17
           High quality sequence stop: 593.
FEATURES   source
           1..598
           /organism="Mus musculus"
           /strain="FVB/N"
           /db_xref="taxon:10090"
           /clone="IMAGE:3969256"
           /clone_lib="NCI_CGAP_Mam1"
           /tissue_type="tumor, biopsy sample"
           /dev_stage="10 months, virgin"
           /lab_host="DH10B"
           /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
           Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
           Library constructed by Life Technologies. Investigator
           providing samples: Gilbert Smith, NIH"
BASE COUNT 141 a 175 c 172 g 110 t
ORIGIN
Query Match 49.7%; Score 344.6; DB 10; Length 598;
Best Local Similarity 93.7%; Pred. No. 3e-57;
Matches 359; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
Qy 195 gctgccggctggagccactgtcacacgcacatggtggcaggcccgctggcaggatcct 254
Db 1 GCTCCGGCTGGAGCCACTGTTCACACGCACATGTCGGCGGGCGCCGTGGCAGGGATCCT 60
Qy 255 ggagcactgcgtgatgtaccctaccatcgactgcgtcaagaccgcgatgcagatctacagcc 314
Db 61 GGAGCATTTGCGTGAATGATACCCGATCGACTGCGTCAAGACCCGGATGCAGAGCTTACAGCC 120
Qy 315 tgaccagctccgcctatcgcaatgtgttgagccctctggaggattataaagacgga 374
Db 121 TGACCCAGCCCGCGCTATCGGAACGTGTGGAGGCTCTCTGGAGAAATCATGAGACAGA 180
Qy 375 gggcctatggagggccctagggggcctgaacgtcacagcaacagcgagggcctgccca 434
Db 181 GGGCCTGTGGAGGCCCATGCGGGGCTGAAGTCAAGTCAAGCAACAGCGCGGGCGCTGCCCA 240
Qy 435 cgccctttatttccctgcgtacgaaaaagttaaaaagacattgagtgatgtaaccaccc 494
Db 241 CGCCCTCTATTTTCCTGCTACGAAAAGTTAAAAAGACATTTAGTGTGACGTAATCCACC 300
Qy 495 tgggggcaatagccatatgccaatggtggccggcgggtgtgtggcaacattactcatga 554
Db 301 AGGGGGCAATAGCCATATTGCCAAATGGTGCAGCCCGATGTGTGGCAGATTACTTCAATGA 360
Qy 555 tgcagccatgaaccctgcggaag 577
||||||||||| |||||||||
```

Db	361	TCGAGCCATGATCCAGCGAAG	363	
----	-----	-----------------------	-----	--

Best Local Similarity	93.1%;	Pred. No. 4.7e-53;	
Matches	338;	Conservative	0; Mismatches 25; Indels 0; Gaps 0;

QY	215	tcaccacgcacatggtggcagcgcgcgtggcaggagatcctgagcactgcgtgatgacc	274
Db	1	TCACCACGCACATGTTGGCGGCGCGCGTGGCAGGATCCTGGAGCATTCGCTGATGTACC	60
QY	275	ccatgcgactgcgtcgaagaccgagatgacagagctctacagcctgacccagctcccgctatc	334
Db	61	CGATCGACTGCGTCAAGACCCGGATGCAGAGCCTACAGCCTGACCCAGCGCCCGCTATC	120
QY	335	gcaatggtgtggaggccctctggaggattataagaacggaggccctatggaggcccatga	394
Db	121	GGAAACGTGTGGAGGCTCTCTGGAGAATCATGAGGACAGAGGCGCTGTGGAGGCGCCATGC	180
QY	395	ggggcctgaacctcacagcaacagcgcagggcctgcccagccctttatttttgcctgct	454
Db	181	GGGGGCTGAACCTCAGAGTAACAGCGCGGGCGCTGCCACGCCCTCTATTTCCTGCT	240
QY	455	acgaaaagttaaaaaacatgagtgcgtgtaatccaccctgggggaatagccatattg	514
Db	241	ACGAAAAGTTAAAAAGACATTTGACTGACGTAATCCACCGGGGCAATAGCCATATTG	300
QY	515	ccaatggtggcggcggtgtgtggcaacattacttcattgcagcgcattggaacccctgcgg	574
Db	301	CCAATGGTCAGCGGATGTGTGGCAGATTACTTCATGATGCAGCCATGAATCCAGCGG	360
QY	575	aaq 577	
Db	361	AAG 363	

RESULT	13
LOCUS	BI045863
DEFINITION	MR3-FN0209-300101-004-h10 FN0209 Homo sapiens cDNA, mRNA sequence.
ACCESSION	BI045863
VERSION	BI045863.1 GI:14452485
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	1 (bases 1 to 480)
AUTHORS	Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., D.H., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE	20202663
COMMENT	Contact: Simpson A.J.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil Tel: +55-11-2704922 Fax: +55-11-2707001 Email: asimpson@ludwig.org.br This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR3&t2=MR3-FN0209-300101-004-h10&t3=2001-01-30&t4=1) Seq primer: puc 18 forward High quality sequence stop: 480.
FEATURES	Location/Qualifiers
source	1..480 /organism="Homo sapiens" /db_xref="taxon:9606" /clone_lib="FN0209"

BI794616	560 bp	mRNA	linear	EST 01-OCT-2001
ic92e02.y1	Melton Normalized Mixed Mouse Pancreas 1 NI-MMS1 Mus musculus cDNA 5' similar to TR:Q23125 Q23125 W02B12.9 PROTEIN. [1]			
BI794616				; mRNA sequence.
BI794616.1	GI:15822341			
EST.				
house mouse.				
Mus musculus				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
1 (bases 1 to 560)				
Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K., Lemishka,I., Scarce,M., Brestelli,J., Gradwohl,G., Clifton,S., Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas ,M., Gibbons M., McCann,R., Cole,R., Tsagarishvili,R., Williams,T., Jackson,Y. and Bowers,Y.				
Endocrine Pancreas Consortium				
Unpublished (2000)				
Other_ESTs: ic92e02.x1				
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue				
Endocrine Pancreas Consortium				
Harvard University, Howard Hughes Medical Institute				
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138				
Tel: 617-495-1812				
Fax: 617-495-8557				
Email: dmelton@biohp.harvard.edu				
Library was constructed by Dr. Douglas Melton DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Juliana Brown (brown@fas.harvard.edu)				
Seq primer: -40RP from Gibco				
High quality sequence stop: 430.				
Location/Qualifiers				
1..560				
/organism="Mus musculus"				
/strain="ICR"				
/db_xref="taxon:10090"				
/clone_lib="Melton Normalized Mixed Mouse Pancreas 1 NI-MMS1"				
/sex="Both for embryonic & newborn, male for adult and adult islet"				
/dev_stages="Embryonic day 10.5, E12.5, E16.5, newborn, adult, mixed"				
/lab_host="DH10B"				
/note="Vector: pSPORT1; Site1: Not I; Site2: Sal I; Five libraries representing E10.5/12.5 pancreatic bud, E16.5 pancreas, newborn pancreas, adult pancreas, and adult islets of Langerhans were seperately constructed using SuperScript Plasmid library kit (Life Technologies). cDNA was made by oligo-dT priming and size-selected by column fractionation. Libraries were amplified once on solid support and plasmid DNA from each library was prepared and mixed in equal amounts. The mixed library DNA was normalized by method #4 from Bonaldo, Lennon, and Soares 1996 Genome Research 6:791-806; 0.5 microgram single-stranded mixed library plasmid DNA was mixed with 5 micrograms PCR product representing mixed library inserts and hybridized to an Ecot of 6. Single-stranded (unhybridized) plasmids were isolated by hydroxyapatite chromatography and used to make this library."				
134 a	158 c	163 g	105 t	
BASE COUNT				
ORIGIN				

Query Match	46.6%;	Score 323;	DB 10;	Length 560;
-------------	--------	------------	--------	-------------

Tel: 302-831-1335
Fax: 302-831-2822
Email: cogburn@udel.edu, www.chickest.udel.edu.

FEATURES	SOURCE
1. High Accuracy: The model achieves a high accuracy rate, indicating its effectiveness in predicting the target variable.	Model Performance Metrics
2. Robustness: The model demonstrates robustness, maintaining performance across different datasets and conditions.	Model Performance Metrics
3. Interpretability: The model's predictions are interpretable, allowing for a clear understanding of the underlying factors influencing the outcome.	Model Performance Metrics
4. Scalability: The model is scalable, capable of handling large volumes of data and complex relationships.	Model Performance Metrics
5. Real-time Processing: The model is designed for real-time processing, enabling immediate decision-making based on incoming data.	Model Performance Metrics
6. Integration with Existing Systems: The model seamlessly integrates with existing data management and analysis systems.	Model Performance Metrics
7. Customizable Parameters: The model offers customizable parameters, allowing for fine-tuning to specific use cases.	Model Performance Metrics
8. Comprehensive Documentation: The model is accompanied by comprehensive documentation, detailing its architecture, usage, and limitations.	Model Performance Metrics
9. Regular Updates: The model is subject to regular updates, ensuring it remains current and effective in the face of changing data patterns.	Model Performance Metrics
10. Support and Maintenance: The model is supported by a dedicated team, ensuring timely assistance and maintenance.	Model Performance Metrics

```

1. 546
/organism="Gallus gallus"
/strain="Commercial broiler and Ottawa Res. Centre
Strains 90 & 21"
/db_xref="taxon:9031"
/clone="pgm2n.pk008.12"
/clone_lib="Normalized Chicken Breast Muscle,
and Epiphyseal Growth Plate cDNA library (pgm2n)"
/sex="Male and Female"
/tissue_type="Breast muscle, leg muscle and epiphyseal
growth plate"
/dev_stage="Breast,leg:Embryo(d19);post-hatch(1d,1,3,5,7,9,
,11 weeks);growth plate(1d,7d,14d post-hatch)"
/lab_host="E. coli EMPH108"
/notes="Vector: pCMVSPORT6; Library made from equivalent
pools of total RNA isolated from each tissue (embryonic
muscle 33.3%; juvenile muscle 33.3%; and epiphyseal growth
plate 33.3% of the final RNA pool). Single pass sequencing
from 5'-end"
197 g 70 t 183 c

```

	BASE COUNT	ORIGIN
1	0.000	0.000
2	0.000	0.000
3	0.000	0.000
4	0.000	0.000
5	0.000	0.000
6	0.000	0.000
7	0.000	0.000
8	0.000	0.000
9	0.000	0.000
10	0.000	0.000
11	0.000	0.000
12	0.000	0.000
13	0.000	0.000
14	0.000	0.000
15	0.000	0.000
16	0.000	0.000
17	0.000	0.000
18	0.000	0.000
19	0.000	0.000
20	0.000	0.000
21	0.000	0.000
22	0.000	0.000
23	0.000	0.000
24	0.000	0.000
25	0.000	0.000
26	0.000	0.000
27	0.000	0.000
28	0.000	0.000
29	0.000	0.000
30	0.000	0.000
31	0.000	0.000
32	0.000	0.000
33	0.000	0.000
34	0.000	0.000
35	0.000	0.000
36	0.000	0.000
37	0.000	0.000
38	0.000	0.000
39	0.000	0.000
40	0.000	0.000
41	0.000	0.000
42	0.000	0.000
43	0.000	0.000
44	0.000	0.000
45	0.000	0.000
46	0.000	0.000
47	0.000	0.000
48	0.000	0.000
49	0.000	0.000
50	0.000	0.000
51	0.000	0.000
52	0.000	0.000
53	0.000	0.000
54	0.000	0.000
55	0.000	0.000
56	0.000	0.000
57	0.000	0.000
58	0.000	0.000
59	0.000	0.000
60	0.000	0.000
61	0.000	0.000
62	0.000	0.000
63	0.000	0.000
64	0.000	0.000
65	0.000	0.000
66	0.000	0.000
67	0.000	0.000
68	0.000	0.000
69	0.000	0.000
70	0.000	0.000
71	0.000	0.000
72	0.000	0.000
73	0.000	0.000
74	0.000	0.000
75	0.000	0.000
76	0.000	0.000
77	0.000	0.000
78	0.000	0.000
79	0.000	0.000
80	0.000	0.000
81	0.000	0.000
82	0.000	0.000
83	0.000	0.000
84	0.000	0.000
85	0.000	0.000
86	0.000	0.000
87	0.000	0.000
88	0.000	0.000
89	0.000	0.000
90	0.000	0.000
91	0.000	0.000
92	0.000	0.000
93	0.000	0.000
94	0.000	0.000
95	0.000	0.000
96	0.000	0.000
97	0.000	0.000
98	0.000	0.000
99	0.000	0.000
100	0.000	0.000

Query Match 43.1%; Score 298.4; DB 10; Length 546;
Best Local Similarity 77.0%; Pred. No. 2.8e-48;
Matches 375; Conservative 0; Mismatches 109; Indels 3; Gaps 1;

	91	gggtggtctgcagcggggcgcttgaggccggggggcccggcggcggggacccggggccctgcagg	150
Qy			
Df	13	GGCCGGCGGCGGGCGCGCGAGGGCTCtGGGCCCGCGGGGAGACTGAGCGGGCTCGG	72
Qy	151	cceccggtacgacaagaatccggactccggccccggactacagagcgctgccgctggagcc	210
Df	73	GGCCCGGAGCGCGCTGC CGC --- CCGCCCCCGATACGAGCGCTGCCGAGGGCGCC	129
Qy	211	actgtcacccgcacatggttggcagggcgctgtaggggatcctggagcactgcgtgatg	270
Df	130	GCCGTGTCCACGCACATGCTTGGCGGGCGCGTGCGGGCATCATGAGGACATTCGCTGTATG	189
Qy	271	taccatacgaactgcgtcaagacccggatgcagagtctacagcctgacctgacccagctgccgc	330
Df	190	TACCCCGTCGATTGCTFCAAGACTCGATGCAGAGCCTGGGCCCGAGCCC GCCGCCGC	249
Qy	331	tatcgcaattgcttgaggccctctbtgaggattataagacggaggcccta tggaggccc	390
Df	250	TACCGCAACCTGCTGGAGGCCCTGTGGCGCATCGTCCGTACCAGGGCGCTGTGGAGGCC	309
Qy	391	atgagggggctgaacttcacagcaacagcgcaaggcctgcccaagcccttattttgc	450
Df	310	ATCGGGGCTGAACATTCACCGCACCGCGGGGGCGGCCACGCCCTCTACTTCGCC	369
Qy	451	tgctacgaaagttaaaaagacattgattgatgaatccacctggggccaatagccat	510
Df	370	TGCTACGAAAAGTTAAAAAGACCGCTTAGGGACGTCATCCACGGGGGGCAATAGCCAT	429
Qy	511	attgccaatggtgcggcggtgtgtggcaaacattactcatgatgcagcccatgaacct	570
Df	430	GTGCGNNCGGTGCACCGGGTGTGTAGCAACATTCCTCCACGACSCACGATGAACCT	489
Qy	571	gcggaaag	577
Df	490	GCAGAAG	496

Search completed: September 28, 2002, 06:22:41
Job time: 9182 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2002, 07:55:42 ; Search time 5287.54 Seconds
(without alignments)
3158.251 Million cell updates/sec

Title: US-09-870-113-7
Perfect score: 798
Sequence: 1 atgcagagtctacagctga.....agtggaggctggcaagtga 798

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pi.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pi.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htgo_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

1	798	100.0	1322	6	AX061229	Sequence
2	798	100.0	1448	9	AF327402	Homo sapi
3	798	100.0	1448	17	HS0303077	Homo sapi
4	576	72.2	1889	9	AF327403	Homo sapi
5	576	72.2	1889	17	HS0303078	Homo sapi
6	520	65.2	1244	9	AF267854	Homo sapi
7	520	65.2	123160	9	AL353719	Human DNA
8	518.4	65.0	1588	9	AK056782	Homo sapi
9	504	63.2	1072	10	BC023172	Mus muscu
10	467.2	58.5	221062	2	AC096351	Rattus no
11	417.8	52.4	483	6	AX061229	Sequence
12	392	49.1	401	6	AX071440	Sequence
13	387.8	48.6	4112	10	AF288621	Mus muscu
14	365.6	45.8	1429	9	AF223466	Homo sapi
15	281	35.2	5869	10	AF361699	Mus muscu
16	276	34.6	1410	9	AF155660	Homo sapi
17	276	34.6	1410	9	AF155660	Homo sapi
18	251.2	31.5	191702	2	AC108878	Mus muscu
19	250	31.3	174445	9	AC051642	Homo sapi
20	250	31.3	182504	2	AC022597	Homo sapi
21	248.4	31.1	108765	2	AF216674	Homo sapi
22	166.8	20.9	818	6	AX012190	Sequence
23	155.6	19.5	434	11	G28167	human STS S
24	153.6	19.2	338	6	AX203265	Sequence
25	150.8	18.9	34305	3	CEW02B12	Caenorhabd
26	141.4	17.7	3037	3	AY060268	Drosophil
27	134.6	16.9	1281	3	OYU45998	Onchocerca
28	125.2	15.7	163205	2	AC099368	Rattus no
29	113.4	14.2	599	6	AX198694	Sequence
30	113.4	14.2	599	6	AX209226	Sequence
31	113.2	14.2	2878	3	AF217402	Drosophil
32	105	13.2	108765	2	AF216674	Homo sapi
33	104	13.0	1243	3	OGU45997	Onchocerca
34	103.4	13.0	858	9	BC015013	Homo sapi
35	91.4	11.5	391	11	G22871	human STS W
36	86	10.8	2096	8	SCMR53G	Yeast nucle
37	86	10.8	2496	8	SCYJL133W	S.cerevisia
38	86	10.8	40724	8	SCX731	S.cerevisia
39	84.8	10.6	500	6	AX336750	Sequence
40	84.8	10.6	657	6	AX011696	Sequence
41	84.6	10.6	1864	8	SCYKR052C	Sequence
42	84.6	10.6	2601	8	SCMR54	Yeast nucle
43	82.2	10.3	76299	2	AC019585	Drosophil
44	82.2	10.3	172363	3	AC009388	Drosophil
45	82.2	10.3	203557	3	AE003763	Drosophil

ALIGNMENTS

RESULT	1					
AX061229	AX061229	Sequence	76 from Patent WO0078953.	DNA	linear	PAT 22-JAN-2001
LOCUS	AX061229	Sequence	76 from Patent WO0078953.	DNA	linear	PAT 22-JAN-2001
DEFINITION	AX061229	Sequence	76 from Patent WO0078953.	DNA	linear	PAT 22-JAN-2001
ACCESSION	AX061229	Sequence	76 from Patent WO0078953.	DNA	linear	PAT 22-JAN-2001
VERSION	AX061229.1	GI:12406365				
KEYWORDS						
SOURCE	human.					
ORGANISM	Homo sapiens					

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1322)
Lal, P., Yang, J., Yue, H., Hillman, J.L., Tang, Y.T., Bandman, O., Burford, N., Baughn, M.R., Azimzai, Y., Lu, D.A., Au-Young, J. and Patterson, C.

Human transport proteins
Patent: WO 0078953-A 76 28-DEC-2000;
Incyte Genomics, Inc. (US)

Location/Qualifiers
1..1322

/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 328 a 376 c 351 g 267 t

ORIGIN		SOURCE	human...
		ORGANISM	Homo sapiens
			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
		REFERENCE	1 (bases 1 to 1448)
		AUTHORS	Li,F.Y., Nikali,K., Grogan,J., Leibiger,I., Leibiger,B., Schweyen,R., Larsson,C. and Suomalainen,A.
		TITLE	Characterization of a novel human putative mitochondrial transporter homologous to the yeast mitochondrial RNA splicing proteins 3 and 4
		JOURNAL	FEBS Lett. 494 (1-2), 79-84 (2001)
		MEDLINE	21195335
		PUBMED	11297739
		REFERENCE	2 (bases 1 to 1448)
		AUTHORS	Li,F., Nikali,K., Grogan,J., Leibiger,I., Leibiger,B., Schweyen,R., Larsson,C. and Suomalainen,A.
		TITLE	Direct Submission
		JOURNAL	Submitted (08-DEC-2000) Human Molecular Genetics, National Public Health Institute, Mannerheimintie 166, Helsinki 00300, Finland
		FEATURES	Location/Qualifiers
		source	1..1448
			/organism="Homo sapiens"
			/db_xref="taxon:9606"
			/chromosome="10"
			/map="10q24"
			30..1124
			/note="similar to mitochondrial RNA splicing protein 3/4; alternatively spliced"
			/codon_start=1
			/product="putative mitochondrial solute carrier splice variant"
			/protein_id="AAK49519.1"
			/db_xref="GI:13926047"
			/translation="MELEGRAGVAGGPAAGPCRPGESALLDGNLQRVGVRGAGGG
			EAGCRPPVRQDPDSDGPDYEALPAGATVTTHMVAGAVGLEHCVMPIDCVKTRMOS
			LQDPDAARYNRVLEALRWRIIRTEGLRPMRLNVTATGAPAHALFYACVEKLKKTLS
			DVTHPGNSHIANGAACVATLLHDAAMNPAEVVKORMQNSPYHRVTPCVRAVWON
			EGAGAFVSVYTTOLTMNVPOAHEMTYFLOEHPORRYNPSHVLSCACAGAVAA
			AATPLDVCKTLTLNTOESLANSHTGHITGMASAFYVQGVGTAYFRGVQARVYI
			QIPSTAIAMSVYEFYLIITKROEFENRAGK"
		BASE COUNT	322 a 408 c 435 g 283 t
		ORIGIN	
		Query Match	100.0%; Score 798; DB 9; Length 1448;
		Best Local Similarity	100.0%; Pred. NO. 7.2e-240;
		Matches 798; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1	atgcagagctctacagcctgacccagctgcccgcctatgcgaatgtgttgaggccctctgg	60
Db	186	ATGCAGAGCTCTACAGCTGACCCAGCTGCCGCTATCCGATGTGTGGAGGCCCTCTGG	245
Qy	61	aggattataagaacgaggggcccctatggaggccctatgaggggctgaacgtcacagcaaca	120
Db	246	AGGATTATAAGAAGGAGGGGCTATGAGGCCCCATGAGGGGCTGAACGTCAAGCAACA	305
Qy	121	ggcagaggcctcccagcccttatttttcctgctcacaaaaagttaaaaaagacattg	180
Db	306	GGCGAGGGCCTGCCAGCCCTTTATTTTCCTGCTACGAAAAGTTAAAAAGACATTG	365
Qy	181	agtatgtaatccaccctgggggcaatagccatatgtccaatggtgcggcggtgtgtg	240
Db	366	AGTGATGTAATCCACCCCTGGGGCAATAGCCATATGCCATGTGTGGCGCGGTGTGTG	425
Qy	241	gcaacattactctatgatgcagccatgaacccctgcggaagtgttcaagcagagatgcag	300
Db	426	GCAACATTACTTTCATGATGCAGCCATGAACCTCGGAAAGTGGTCAAGCAGAGGATCGAG	485
Qy	301	atgtacaactcacataccaccgggtgacagactgtgtacgggcagtggtgcaaaatgaa	360
Db	486	ATGTACAACACTCACCATACCACCGGGTGACAGACTGTGTACGGGCAGTGTGGCAAAATGAA	545
Qy	361	ggggccggggcctttaccgcagctcacacccagctgcacatgaacgttcctttccaa	420
Db	546	GGGGCCGGGGCTTTTACCGCAGCTACACACCAGCTGACCATGAACGTTCCTTTCCAA	605
Qy	421	gccattcaactcatgaactatgaattcctgcagagcaactttaacccccagagcgttac	480
Db	606	GCCATTCACTTCATGACCTATGAATTCCTGCAGGAGCACITTTAACCCCCAGAGACGGTAC	665
Qy	481	aacccaagctcccagctcctctctgagcttgagctgcagagagctgtagctgcggagccaca	540
Db	666	AACCCAAGCTCCCAAGCTCTCTCGAGCTTGCAGAGAGCTGTAGTGGCGGAGCCACA	725
Qy	541	accacactggagctttgcaaacactgctcaacccagagctcttggctttgaactca	600
Db	726	ACCCACTGGAGCTTGCAAAACACTGCTCAACACCAGAGTCTTGGCTTTGAATCA	785
Qy	601	cacattacaggacatacacaggcatggctagtgcccttcaggacgggtatatcaagtaggt	660
Db	786	CACATTACAGGACATATCACAGGATGGCTAGTGCCTTCAGGACGGTATATCAAGTAGGT	845
Qy	661	gggtgaccgcctatttccaggggtgaggccagagataattaccagatccctccaca	720
Db	846	GGGTGACCGCCTATTTCCGAGGGGTGCAGGCCAGAGTAATTTACAGATCCCTCCACA	905
Qy	721	gccatcgatggtctgtatgagttcttcaaatcacctaatcacataaaaggcaagaagag	780
Db	906	GCCATCGATGGTCTGTGTATGAGTTCTTCAATACCTAATCACTAATAAGCAGAGAG	965
Qy	781	tggagggtggcaagtga 798	
Db	966	TGGAGGCTGGCAAGTGA 983	
RESULT 2			
AF327402			
LOCUS	AF327402	1448 bp	mRNA linear PRI 02-MAY-2001
DEFINITION	Homo sapiens putative mitochondrial solute carrier splice variant mRNA, complete cds, alternatively spliced, nuclear gene for mitochondrial product.		
ACCESSION	AF327402		
VERSION	AF327402.1	GI:13926046	
KEYWORDS			

FT YLITRQEWBRAGK"

XX Sequence 1889 BP; 404 A; 508 C; 544 G; 433 T; 0 other;

SQ Query Match 72.2%; Score 576; DB 17; Length 1889; Best Local Similarity 100.0%; Pred. No. 6.3e-170; Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 223 ggtgcggcggtgtgtggaacattacttacttcatgatgcagccatgaacccctcggaagtg 282
DB 990 GTTGGCGCGGGTGTGTGCACATTACTTTCATGATGACGCCATGAACCTCGGGAAGTG 1049
QY 283 gtcaagcagagatgcagatgtacaactcaccatcaccacccgggtgacagactgtgtacgg 342
DB 1050 GTCAAGCAGAGATGCAGATGTACAACATCACCATACACCGGTGACAGACTGTGTACGG 1109
QY 343 gcagttgtgcaaaatgaagggccggggcctttaccgcagctacacacccagctgaac 402
DB 1110 GCAGTGTGCAAAATGAAGGGCCGGGGCTTTTACCGCAGCTACACCAACCGCTGACC 1169
QY 403 atgaacgttctcttccaaagccattcacttcacatgacctatgaattcctgcagagcacttt 462
DB 1170 ATGAAGTTCTTTCACAGCCATTCACTTCATGACCTATGAATTCCTGCAGGAGCACATT 1229
QY 463 aacccccagagcgtgtacaacccaagctcccccgtctctctgtgagcttgcgagagct 522
DB 1230 AACCCCCAGAGACGGTACAAACCAAGCTCCACAGCTCTCTGGAGCTTGGCGAGGAGCT 1289
QY 523 gtacgtgcgcagccacacaccccaactggacgtttgcacaaacactgcctcaacacccaggag 582
DB 1290 GTAGCTGCCGAGCCACAAACCCCACTGGACGTTTGCAAAACCACTGCTCAACACCCAGGAG 1349
QY 583 tcccttggtttgaactcacacattacagacacatcacagggatgctgtagtgccttcag 642
DB 1350 TCCTTGGCTTTGAACCTCACACATTACAGGACATATCACAGGATGGCTAGTGCCTTCAGG 1409
QY 643 acggtatatcaagttagtggtggtgacccgctattcccgaggggtgcagccagagtaatt 702
DB 1410 ACGGTATATCAAGTAGTGGGTGACCGGCTATTTCCGAGGGGTGCAGGCCAGAGTAATT 1469
QY 703 taccagatccctccacagccatgcagctgctgtgtatgaatttccaaatacctaattc 762
DB 1470 TACCAATCCCCCTCACAGCCATCGCATGGTCTGTGTATGAGTTCTTCAATACCTAATC 1529
QY 763 actaaaggcaagaagtgaggggctggcaagtga 798
DB 1530 ACTAANAGCAGAGAGTGGAGGGCTGGCAAGTGA 1565

RESULT 6 AF267854 1244 bp mRNA linear PRI 02-JAN-2001
LOCUS Homo sapiens NP0016 mRNA, complete cds.

DEFINITION AF267854
ACCESSION AF267854.1 GI:12006034
VERSION AF267854.1
KEYWORDS FLI_CDNA.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE Xu,X., Yang,X., Gao,G., Xiao,H., Chen,Z. and Han,Z.
AUTHORS 1 (bases 1 to 1244)
TITLE Direct Submission
JOURNAL Submitted (25-MAY-2000) Chinese National Human Genome Center at Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, People's Republic of China
Location/Qualifiers
FEATURES
source 1..1244
/organism="Homo sapiens"
/db_xref="taxon:9606"

/tissue_type="normal pituitary"
381..911
CDS

/codon_start=1
/product="NP0016"
/protein_id="AAG44723.1"
/db_xref="GI:12006034"

/translation="MALLVYKQRMQYNSPHRYTDCVRAYWQNEGAGAFYRSYTTQL
TMNVFFQAIHFMTEYFQEHFNPPORRYNSHVLGSGACAGAVAAAATPLDVKLTLLN
TOESLALNSHITGHITGNASAFYVQVGGVTAYFRGVOARVIYIIPSTAWSVYEF
FKYLITRQEWBRAGK"

BASE COUNT 308 a 331 c 290 g 315 t
ORIGIN

Query Match 65.2%; Score 520; DB 9; Length 1244;
Best Local Similarity 100.0%; Pred. No. 2.7e-152;
Matches 520; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 aqgtgtcaagcagaagatgcagatgtacaactcaccatcaccacccgggtgacagactgtgt 338
DB 392 AGTGTCTAAGCAGAGATGTCAGATGTACAACATCACCATACACCGGGTGACAGACTGTGT 451
QY 339 acgggcagctgtgtgcaaaatgaagggccggggccttttaccgcagctacacacccagct 398
DB 452 ACGGGCAGTGTGGCAAAATGAAGGGCCGGGGCCCTTTTACCACAGCTACACACCCAGCT 511
QY 399 gacctgaacgttctcttccaaagccattcacttcacatgacctatgaattcctgcaggagca 458
DB 512 GACCATGAACGTTCTCTTCCAAAGCCATTCACTTCATGACCTATGAATTCCTGCAGGAGCA 571
QY 459 ctttaacccccagagcgtgtacaacccaagctcccaagctctctctctgtgagcttgcaggg 518
DB 572 CTTTAACCCCCAGAGACGGTACAAACCAAGCTCCACAGCTCTCTGTGAGCTTGGCAGG 631
QY 519 agctgtagctgcgcagccacacaccccaactggacgttttgcacaaacactgctcaacaccca 578
DB 632 AGCTGTAGCTGCGCAGCCACACACCCCACTGGACGTTTGCAAAACACTGCTCAACACCCA 691
QY 579 gtagtctcttggttgaactcacacattacagagacatcacagggcatgagctagtcctt 638
DB 692 GGAGTCTTGGCTTTGAACTCACACATTACAGGACATATCACAGGCATGGCTAGTGCCTT 751
QY 639 caggcaggtatataagtagtggtgggtgacgccttatitccgaggggtgcagggcagagct 698
DB 752 CAGGACGGTATATCAAGTAGGTGGGGTGACCGGCTATTTCCGAGGGGTGCAGGCCAGAGT 811
QY 699 aattaccagatccctccacagccatcgactgctgtgtatgagttcttcaataacct 758
DB 812 AATTTACCAGATCCCTCCACAGCCATCGCATGGTCTGTGTATGAGTTCCTTCAATACT 871
QY 759 aatcactaaaggcaagaagatggagggtggcaagtga 798
DB 872 AATCACTAAAGGCAAGAAGAGTGGAGGGCTGGCAAGTGA 911

RESULT 7
LOCUS AL353719/c

DEFINITION Human DNA sequence from clone RP11-85A1 on chromosome 10, complete sequence.
ACCESSION AL353719 AC007643
VERSION AL353719.10 GI:15787725
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 123160)
AUTHORS Direct Submission
TITLE Ramsay,H.
JOURNAL Submitted (25-SEP-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Sep 26, 2001 this sequence version replaced gi:14280413.
During sequence assembly data is compared from overlapping clones.

Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em', EMBL; Sw', SWISSPROT; Tr', TREMBL; Wp', WORMPEP; Information on the WORMPEP database can be found at

<http://www.sanger.ac.uk/Projects/C.elegans/wormpep> This sequence

was generated from part of bacterial clone contigs of human

chromosome 10, constructed by the Sanger Centre Chromosome 10

Mapping Group. Further information can be found at

<http://www.sanger.ac.uk/HGP/Chr10>

RP11-85A1 is from the library RPI-11.1 constructed by the group of

Pieter de Jong. For further details see

<http://www.ncbi.nlm.nih.gov/bacpac/home.htm>

VECTOR: pBACE3.6

IMPORTANT: This sequence is not the entire insert of clone

RP11-85A1. It may be shorter because we sequence overlapping

sections only once, except for a short overlap.

The true left end of clone RP11-85A1 is at 1 in this sequence. The

true left end of clone RP11-483F11 is at 123061 in this sequence.

The true right end of clone RP11-129J12 is at 51589 in this

sequence.

```

FEATURES             Location/Qualifiers
     source            1..123160
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /chromosome="10"
                        /clone="RP11-85A1"
                        /clone_lib="RPI-11.1"
                        /notes="Single clone region. Sequence from reads from a
                        short insert library derived from a single pUC clone.
                        Restriction digest data confirm the assembly."
     misc_feature      complement(7066..7118)
     unsure            52162
     unsure            52167
     BASE COUNT        33221 a 28084 c 27196 g 34659 t
     ORIGIN
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2210
2211
2212
2213
2214
2215
2216
2217
2218
2219
2220
2221
2222
2223
2224
2225
2226
2227
2228
2229
2230
2231
2232
2233
2234
2235
2236
2237
2238
2239
2240
2241
2242
2243
2244
2245
2246
2247
2248
2249
2250
2251
2252
2253
2254
2255
2256
2257
2258
2259
2260
2261
2262
2263
2264
2265
2266
2267
2268
2269
2270
2271
2272
2273
2274
2275
2276
2277
2278
2279
2280
2281
2282
2283
2284
2285
2286
2287
2288
2289
2290
2291
2292
2293
2294
2295
2296
2297
2298
2299
2300
2301
2302
2303
2304
2305
2306
2307
2308
2309
2310
2311
2312
2313
2314
2315
2316
2317
2318
2319
2320
2321
2322
2323
2324
2325
2326
2327
2328
2329
2330
2331
2332
2333
2334
2335
2336
2337
2338
2339
2340
2341
2342
2343
2344
2345
2346
2347
2348
2349
2350
2351
2352
2353
2354
2355
2356
2357
2358
2359
2360
2361
2362
2363
2364
2365
2366
2367
2368
2369
2370
2371
2372
2373
2374
2375
2376
2377
2378
2379
2380
2381
2382
2383
2384
2385
2386
2387
2388
2389
2390
2391
2392
2393
2394
2395
2
```

Db	745	AGTGGTCAAGCAGAGATGAGATGACAACTACACATACCACCGGTGACAGGCTGTGT	
QY	339	acgggcagtggtgcaaatgaaggccggccctttaccgcagctacacccacgct	
Db	805	ACGGGCAGTGTGGCAAAATGAAGGGCCGGGGCTTTTACCGCAGCTACACCCAGCT	
QY	399	gacctgaacgtctcttccaagccattcacttcacatcatgaattcctgcagagca	
Db	865	GACCATGAACGTTCTCTTCCAAAGCCATTACATTTCATGACCTATGAATTCTCTGCAGGAGCA	
QY	459	ctttaacccccagagcgtacaaacccaagctcccaagctcctctctgagcttgccag	
Db	925	CTTTAACCCTCCAGAGCGGTACAACCAAGCTCCCAAGCTCCCTCTCTGGAGCTTGGCAGG	
QY	519	agctgtagctgcccagcagcacacccccactggacgtttgcaaaacactctcaaaccca	
Db	985	AGCTGTAGTGGCGCAGCCACAAACCCACTGGACGTTTGCARAAACTGCTCAACACCCA	
QY	579	ggagtccttggcttgaactcacattacagagacatatcacaggcagcatggctagtcctt	
Db	1045	GGAGTCTTGGCTTTTGAACATCACATTACAGGACATATCACAGGCATGGCTAGTGCCTT	
QY	639	caggacggtatatcaagtagtggtgggtgaccgcctatttccagggggtgcagggcagag	
Db	1105	CAGGACGGTATATCAAGTAGTGGGTGACCGCCTATTTCAGAGGGGTGCAGGCCAGAGT	
QY	699	aatttaccagatccctccacagcaccatcgatggctgtgtatgattcttcaatacct	
Db	1165	AATTTACAGATCCCTCCACAGCCATCGATGGTCTGTGTATGAGTTCTTCAATAACCT	
QY	759	aatcactaaaaggcaagaagtgagggtggtgcgaagtga	
Db	1225	AATCACTAAAGGCAAGAGATGGAGGGCTGGCAAGTGA	
RESULT	9		
LOCUS	BC023172		
DEFINITION	Mus musculus, 1072 bp mRNA linear ROD 07-FEB-2002		
ACCESSION	clone MGC:37028 IMAGE:4949779, mRNA, complete cds.		
VERSION	BC023172		
KEYWORDS	MGC.		
SOURCE	BC023172.1 GI:18606247		
ORGANISM	house mouse.		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
TITLE	1 (bases 1 to 1072)		
JOURNAL	Strausberg, R.		
REMARK	Submitted (04-FEB-2002) National Institutes of Health, Mammalian		
COMMENT	Gene Collection (MGC), Cancer Genomics Office, National Cancer		
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,		
	USA		
	NIH-MGC Project URL: http://mgc.nci.nih.gov		
	Contact: MGC help desk		
	Email: cgraps@mail.nih.gov		
	Tissue Procurement: Jeffrey Green M.D.		
	cDNA Library Preparation: Life Technologies, Inc.		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: Baylor College of Medicine Human Genome		
	Sequencing Center		
	Center code: BCM-HGSC		
	Web site: http://www.hgsc.bcm.tmc.edu/cdna/		
	Contact: amg@bcm.tmc.edu		
	Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,		
	Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,		
	Richards, S., Gibbs, R.A.		
	Clone distribution: MGC clone distribution information can be found		
	through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov		
	Series: IRAK Plate: 60 Row: h Column: 15		

This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.		
FEATURES	Location/Qualifiers	
source	1..1072	
	/organism="Mus musculus"	
	/db_xref="taxon:10090"	
	/clone="MGC:37028 IMAGE:4949779"	
	/tissue_type="Mammary tumor, C3(1)-Tag model. Infiltrating ductal carcinoma, 5 month old virgin mouse."	
	/clone_lib="NCI_CGAP_Mam6"	
	/lab_host="DH10B"	
	/note="Vector: pCMV-SPORT6"	
CDS	188..721	
	/codon_start=1	
	/product="Similar to putative mitochondrial solute carrier"	
	/protein_id="AAH23172.1"	
	/db_xref="GI:18606248"	
	/translation="MNPAAEVKQRMQYNSPYHRVDCVRAVWQNEGAGAFRSYTTQ LTMVYFOAIHFMTYFLOEHFNPPORRYNPSSHVLGCAGACAGAVADAATPLDVKCKLL NTOESIALNSNITGHITGMASAFRTVYQVGGTVAFRGVQARVIYIIPSTAIASVWE PFKYLITKRQEEWRAGK"	
BASE COUNT	268 a 308 c 263 g 233 t	
ORIGIN		
Query Match	63.2%;	Score 504; DB 10; Length 1072;
Best Local Similarity	92.2%;	Pred. No. 2.9e-147;
Matches	531; Conservative 0;	Mismatches 45; Indels 0; Gaps 0;
QY	223	ggtgcggccgggtgtgtggaacattacttcattgatgcagccatgaacccctgcggaagtg 282
Db	146	GTGTCAGCCGGATGTGTGGCGACATTACTTCATGATGCAGCCATGAATCCACGGGAAGTC 205
QY	283	gtcaagcagagatgcagatgtacaactcaccatccaccgggtgcagactgtgtcagg 342
Db	206	GTCAAGCAGAGATGCAGATGTACAACTCGCGTACCACCGCTGCAGACTGTGTTCGG 265
QY	343	gcagtggtgcaaaaagaaggccggccgtttaccgcagctacacacccagctgacc 402
Db	266	GCAGTGTGGCAAAATGAAGGGCCGGGGCTTTTACCAGGCTACACGACCCAGCTGACC 325
QY	403	atgaacgttctcttccaagccattcacttcacatgaacctatgaattctctgcagagcacatt 462
Db	326	ATGATGTCCCTTCCAAGCCATTCACTTCATGACCTATGAGTTCTTGCAGAGACACTTT 385
QY	463	aacccccagagcgtgtacaaccccaagctccccacgtctctctgtgagcttgcgcaggagct 522
Db	386	AACCCCCAGAGCGGTACAACCCAGCTCCCATGTGCTGTGTGGGCGCTGCCAGGAGCA 445
QY	523	gtagctgcgcagcagcacacacccactggacgtttgcacaaactgcctaaacacccaggag 582
Db	446	GTAGCTGACGCGCCGCACAAACCCACTGGACGTTTGTAAACACACTGCTCAACACCCAGGAA 505
QY	583	tccttggctttgaactcacacattacagacacatacacagcagcatgctagtccttcagg 642
Db	506	TCCCTGGCTTTGAACCTCAACATTACTGTGACACATCAACAGGCATGCTAGTGCCCTTCAGG 565
QY	643	acggtatatcaagtaggtgggtgacccgctatttccagggggtgcagccagagtaatt 702
Db	566	ACGGTCTATCAAGTAGCGGGGTGACTGCTTACTTCCGAGGGGTGCAGGCCAGAGTCATT 625
QY	703	taccagatccccctccacagccatcgcagtgtgtgtatgattcttcttaataacctaattc 762
Db	626	TACCATGTCCTCCACAGCCATCGCATGGTGTGTGTATGAATTTTCAATACCTAATC 685
QY	763	actaaagcgaagaagtagtgagggtgcgaagtga 798
Db	686	ACAAACGGCAAGGAATGGAGGGCAGGCAAGTGA 721
RESULT	10	

AC096351/c
LOCUS
DEFINITION
Rattus norvegicus chromosome Rf1 clone CH230-24M6, WORKING DRAFT
SEQUENCE, 33 unordered pieces.
ACCESSION
AC096351
VERSION
GI:17944054
KEYWORDS
HTG: HTGS_PHASE1; HTGS_DRAFT.
SOURCE
Norway rat.
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 221062)
Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C.,
Albrooks, S.L., Amaral, H.C., Are, J.R., Banks, T., Barbarella, J.,
Benton, J., Bimaga, K., Blankenburg, K., Bonnin, D., Bouck, J.,
Bowles, S., Brieve, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C.,
Burch, P., Burckett, C., Burrell, K.L., Byrd, N.C., Carron, T.F.,
Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R.,
Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C.,
Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C.,
Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O.,
Denn, A.L., Ding, X., Dinh, H.H., Douthwaite, K.J., Draper, H.,
Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C.,
Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J.,
Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T.,
Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S.,
Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A.,
Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C.,
Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J.,
Joudah, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S.,
Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J.,
Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C.,
Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W.,
Loulsegh, H., Lozada, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R.,
Mar, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A.,
Martinez, E., Massey, E., Mahoney, E., McLeod, M.P., Meador, M.,
Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K.,
Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N.,
Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S.,
Ogih, M., Okwuonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B.,
Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L.,
Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Rolfe, M.,
Ruiz, S., Savary, G., Scherer, S., Scott, G., Shen, H., Shoohtari, N.,
Sisson, I., Sodergren, E., Sonaite, T., Sparks, A., Stanley, H.,
Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K.,
Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, R.,
Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R.,
Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C.,
Watlington, S., Williams, G., Williamson, A., Wleczyk, R., Wooden, S.,
Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.
Direct Submission
Unpublished
2 (bases 1 to 221062)
Worley, K.C.
Direct Submission
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Dec 20, 2001 this sequence version replaced gi:15627972.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GE2D
Center clone name: CH230-24M6
----- Summary Statistics
Assembly program: Phrap; version 0.990329First call to
findPhrap.pl
Consensus quality: 200164 bases at least Q40

Consensus quality: 205303 bases at least Q30
Consensus quality: 209366 bases at least Q20
Estimated insert size: 209880; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-ef estimation
Quality coverage: 3.9x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 33 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved
* 1
* 32851: contig of 32851 bp in length
* 32852 32951: gap of unknown length
* 32952 49984: contig of 16033 bp in length
* 49985 49984: gap of unknown length
* 63815: contig of 14731 bp in length
* 63915: gap of unknown length
* 63916 76005: contig of 12090 bp in length
* 76006 76105: gap of unknown length
* 76106 88842 88941: gap of unknown length
* 88842 100591: contig of 11650 bp in length
* 100592 100691: gap of unknown length
* 100692 111702: contig of 11011 bp in length
* 111703 111802: gap of unknown length
* 111803 122664: contig of 10862 bp in length
* 122665 122764: gap of unknown length
* 122765 129711: contig of 6947 bp in length
* 129712 129811: gap of unknown length
* 129812 132447: contig of 7436 bp in length
* 132448 137347: gap of unknown length
* 137348 144676: contig of 7329 bp in length
* 144677 144776: gap of unknown length
* 144777 151224: contig of 7348 bp in length
* 151225 152224: gap of unknown length
* 152225 158878: contig of 6654 bp in length
* 158879 158978: gap of unknown length
* 158979 163928: contig of 4950 bp in length
* 163929 164028: gap of unknown length
* 164029 169353: contig of 5325 bp in length
* 169354 169453: gap of unknown length
* 169454 175590: contig of 6137 bp in length
* 175591 175690: gap of unknown length
* 175691 180475: contig of 4785 bp in length
* 180476 180575: gap of unknown length
* 180576 184657: contig of 4082 bp in length
* 184658 184757: gap of unknown length
* 184758 190332: contig of 5575 bp in length
* 190333 190432: gap of unknown length
* 190433 194134: contig of 3702 bp in length
* 194135 194234: gap of unknown length
* 194235 196592: contig of 2358 bp in length
* 196593 196692: gap of unknown length
* 196693 200530: contig of 3838 bp in length
* 200531 200630: gap of unknown length
* 200631 202809: contig of 2179 bp in length
* 202810 202909: gap of unknown length
* 202910 205034: contig of 2125 bp in length
* 205035 205134: gap of unknown length
* 205135 207176: contig of 2042 bp in length
* 207177 207276: gap of unknown length
* 207277 208845: contig of 1569 bp in length
* 208846 208945: gap of unknown length
* 208946 211271: contig of 2326 bp in length
* 211272 211371: gap of unknown length
* 211372 213302: contig of 1931 bp in length
* 213303 213402: gap of unknown length
* 213403 215484: contig of 2082 bp in length

BASE COUNT	100 a	104 c	120 a	76 t	1 others
BASE COUNT	100 a	104 c	120 a	76 t	1 others

ORIGIN

Query Match 49.1%; Score 392; DB 6; Length 401;
Best Local Similarity 100.0%; Pred. No. 5.2e-112; Indels 0; Gaps 0;
Matches 392; Conservative 0; Mismatches 0;

Qy 8 gtctacacgctgacccagctgcccgcctatcgcaatgtgttgaggccctctggaggatta 67
Db 9 gTCTACAGCCTGACCCAGCTGCCCGCTATCGCAATGTCTTGGAGGCCCTCTGGAGGATTA 68
Qy 68 taagAACGAGGCGCTATGAGGCGCCATGAGGGGCTGAACGCTCACAGCAACAGCGCGAG 127
Db 69 TAAGAACGAGGCGCTATGAGGCGCCATGAGGGGCTGAACGCTCACAGCAACAGCGCGAG 128
Qy 128 ggcctgcccacgccccttatttgcctctactagaaaaagttaaaaaagacattgagtgtg 187
Db 129 GGCCTGCCACGCCCTTTATTTTGCCTGCTACGAAAGTTTAAAAAGACATTTGAGTGATG 188
Qy 188 taatccacccctgggggcaatagccatattgcaatgtgtgcccgggtgtgtgccaacat 247
Db 189 TAATCCACCCCTGGGGCAATAGCCATATTGCCAATGTGTGGCCGGGTCTGTGGCAACAT 248
Qy 248 tacttcattgacgagccatgaacccctcggaagtgtgtcaagcagaggtgacagatgtaca 307
Db 249 TACTTTCATGATGACCCATGAACCCCTGCGGAAGTGTGCAAGCAGAGGATGCGATGTACA 308
Qy 308 actcaccataccacgggtgacagactgtgtacggcagctgtgcaaaataaaggggccg 367
Db 309 ACTCACCATACACCGGTGACAGACTGTGTACGGGCAGTGTGGCAAAATGAAGGGGCCG 368
Qy 368 gggccttttacgcagctacacacccagctg 399
Db 369 GGGCCTTTTACCGCAGCTACACCACCCAGCTG 400

RESULT 13

AF288621 4112 bp mRNA linear ROD 27-OCT-2001
LOCUS Mus musculus mitochondrial solute carrier-like protein mRNA,
complete cds.

ACCESSION AF288621
VERSION AF288621.1 GI:16506177

KEYWORDS house mouse.
SOURCE Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Li,Q.-Z., Ruan,Q.-G., Eckenrode,S., Shi,J.-D., Cruz,P., Wang,C.-Y.
and She,J.-X.

TITLE A new gene which is highly expressed in NOD mice spleen

JOURNAL Unpublished

REFERENCE 1 (bases 1 to 4112)

AUTHORS Li,Q.-Z., Shi,J.-D., Cruz,P., Echenrode,S., Ruan,Q.-G. and Wang,C.-Y.

TITLE Direct Submission

JOURNAL Submitted (20-JUL-2000) Pathology, University of Florida, 1600 SW

Archer Road, Rm. D6-15, Gainesville, FL 32610, USA

FEATURES

source

1. 4112

/organism="Mus musculus"

/db_xref="taxon:10090"

116..1132

/note="50E12"

/codon_start=1

/product="mitochondrial solute carrier-like protein"

/protein_id="AA123859.1"

/db_xref="GI:16506178"

/translation="MELRRGVGNQAGRWGDCRDGGCGSKDAGSEYENLPISAS

VSTHMTACAMAGILEHSIMTPVDSVKTRMOSLPDPAKRTISYIGALKRIMHTEGWFR

PLRLGNVMMGAGPAHMYFACYENMKRTLNDVFSHGNSHLANGVAGSMATLLHDAY

MNPAEYVVKQLQYNSHQSAFSCIIRTVWRTEGLGAFYRSYTTQLTMNIPFQSIHFI

YEFLOEQVNPDRDYNPOSHIISGLAGALAAATAATPLDVCKTLINTQENMALSLANVS
GRISGMANAFRTVYQLNGLAGYFKGIQARVIYQMPSTAIWSVEFFKYLITRKOLEN
RTLV"

BASE COUNT 970 a 1024 c 970 g 1148 t
ORIGIN

Query Match 48.6%; Score 387.8; DB 10; Length 4112;
Best Local Similarity 69.3%; Pred. No. 1.3e-110;
Matches 543; Conservative 0; Mismatches 237; Indels 3; Gaps 1;

Qy 1 atgcagagcttcacagctgaccagctgcccgcctatcgcaatgtgttgaggccctctgg 60
Db 332 ATGCAGAGCTTTGAATCCAGATCCCAAGCCCGGTATACAAGCATCTATGCGCCCTCAAG 391
Qy 61 aggtattataaagaacgggggctctatgagggccatgagggggtgaacgtcacagcaaca 120
Db 392 AGGATCATGCACACTGAAGGCTTCTGGAGGCCCTTGGGGCCCTGAACGTGATGATG 451
Qy 121 ggcgaggggctgcccacgccccttatttgcctgctacgaaagttaaaaaagacattg 180
Db 452 GGTGCAGGGGCCCGCCACGCCCATGTATTTTGCCTGCTATGAAAACATGAAAAGACTTTA 511
Qy 181 agtgcgttaattccacctgggggcaatagccatattgccaatgtgtgcggcggtgtgtg 240
Db 512 AATGACGTTTTTACGCCACCAAGGAAACAGCCATCTAGCTAATGGGTAGCTGGGAGTATG 571
Qy 241 gcaacattacttcattgatgcagccatgaacccctcggaagtgtgcaagcagagagtcag 300
Db 572 GCCACCCCTACTCCACAGTGCAGTAAATGAATCCAGCAGAAGTGGTGAACACGCGTTAC 631
Qy 301 atgtacaactcaccataccacccgggtgacagactgtgtacgggcagctgtgcaaaatgaa 360
Db 632 ATGTACAACTCCAGACCCAGCTAGCCCTTCAGTTGTATCCGGCAGAGTGTGGCGGACCG 691
Qy 361 ggggcccgggcccctttaccgcagctacacccagctgaccatgaacgttccttccaa 420
Db 692 GGGTTGGGGGCTTCTACAGGAGTTACACACACAGCTGACCATGAATATCCCTTCCAG 751
Qy 421 gccattcaattcattgacattatcctgcagagacacttaacccccagagacggtac 480
Db 752 TCAATTCACTTCACTCAGCTATGAGTTTTCGAGGAGCAAGTCAACCCCTCGCGGGACTAC 811
Qy 481 aacccaagctccacgctcctctctgtgagcttgcgagagctgtagctgcccgcagccaca 540
Db 812 AACCCAGAGTCTACATCATCTCAGGAGGCTTGGCGGGGCATGCGCGAGCTGCCACC 871
Qy 541 accccactggagctttgcacaaacactgctcaacccccaggagctgctgttgaac--- 597
Db 872 ACCCGCTGGAGCTGTGCAAAACCCCTCTCAACACGAGGAGAACATGGCTCTCTCCCTG 931
Qy 598 tcacacattacaggacatatcacaggcatggctagctgaccttcaggagcggtatataagta 657
Db 932 GCCAACCTACAGCGCCGCGCTGTGCGGCATGGCCAAATGCTTCCGGACGGTGTACAGCTC 991
Qy 658 ggtggggtgacgcctatttccgaggggtgcaggccagagtaatttaccagatccccccc 717
Db 992 AACGGCCTTGGCGGCTATTTCAAAGGCATCCAGGCTCGAGTCATCTACCAGATGCCCTTC 1051
Qy 718 acagccatcgcatggtctgtgtatgagttcttcaaatcacctaatcacctaaaagcaaga 777
Db 1052 ACCGCCATCTCCTGGTCGCTTATGAGTTCTTCAAGTACATCTCTTACAAAGAGGCGCTG 1111
Qy 778 gag 780
Db 1112 GAG 1114

RESULT 14

AF223466

LOCUS

DEFINITION

ACCESSION

AF223466 1429 bp mRNA linear PRI 17-APR-2000
Homo sapiens HT015 protein (HT015) mRNA, complete cds.

VERSION	AF223466.1	GI:7578782	
KEYWORDS	human.		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 1429)		
AUTHORS	Song, H., Gao, G., Peng, Y., Ren, S., Chen, Z. and Han, Z.		
TITLE	A novel gene expressed in human hypothalamus		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 1429)		
AUTHORS	Song, H., Gao, G., Peng, Y., Ren, S., Chen, Z. and Han, Z.		
TITLE	Direct Submission		
JOURNAL	Submitted (12-JAN-2000) Chinese National Human Genome Center at Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, China		
FEATURES	Location/Qualifiers		
source	1..1429		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
gene	1..1429		
	/tissue_type="hypothalamus"		
CDS	/gene="Ht015"		
	45..1088		
	/gene="Ht015"		
	/codon_start=1		
	/product="Ht015 protein"		
	/protein_id="AAF64141.1"		
	/db_xref="GI:7578783"		
	/translation="MELRSGVGSQAVARRMDGSDRGSGGKDATGSEYENLPTSAS VSTMTAGMAGILEHSIMYPVDSVKTRMQLNDVPHOGNSHLANGIAGSMATLLHDAY PLRGVNMVIMGAPAHAMYPACVENMKRTLNDVPHOGNSHLANGIAGSMATLLHDAY MNPFAVYKRLQYNSQHSRAISCIPTWRTGEGAFYRSTYTLTQNIIPQSIHFT YEFLEQVNPHTYNPQSHIISGLGALAGAAAATPLDVCKTLINTQENVALSIANT GRLSGMANARTVYQLNGLPATSKASRRVSSTRCPPLFLGLSMSSSSTFSPASWKI ELHTKGRDHIFS"		
BASE COUNT	326 a 426 c 402 g 275 t		
ORIGIN			
Query Match	45.8%;	Score 365.6;	DB 9; Length 1429;
Best Local Similarity	68.7%;	Pred. No. 1.2e-103;	
Matches	533;	Conservative 0;	Mismatches 239; Indels 4; Gaps 2;
Qy	1	atgcagagtctacagcctgaccagctgcccgtatgctgcaaatgtgttgaggccctctgg	60
Db	261	ATGCAGAGTTTGAGTCCAGATCCCAAGCCAGTACACAACTATCTACGGAGCCCTCAAG	320
Qy	61	aggattataagaacgaggcctatgaggcccatgagggcctgagggcctgaacgtcacagcaaca	120
Db	321	AAAATCATCGGACCGAAGGCTTCTGGAGGCCCTTGGAGGCGCTCAAGTCATGATCATG	380
Qy	121	ggcagaggcctcccagcccttatttgcctgtctacaaaaagttaaaaaaacattg	180
Db	381	GGTCAGGCGCGCCGACATGATTTTCCCTGCTATGAAACATGAAAGGACTTTA	440
Qy	181	agtgatgtaatccaccctgggggcaatagccatatgtgccaatggtgcggcgggtgtgtg	240
Db	441	AATGACGTTTTCCACCACCAAGAAACAGCCACCTAGCCACGGATAGCTGGGAGTATG	500
Qy	241	gcacattactctatgatgagccatgaacccctcggaagtgggtcaagcagagatgcag	300
Db	501	GCCACCTGCTCCACGATGCGGTAAATGAATCCAGAGAAGTGGTGAAGCAGCGCTTGCA	560
Qy	301	atgtacaactacattaccaccgggtacacagctgtatcagggcaggtggcacaatgaa	360
Db	561	ATGTACAACTCGACACCGGCTGACGAATCAGCTGCAATCCGCGGCTGGAGGACCGAG	620
Qy	361	ggggccgggcccctttaccgcagctacaccaccagctgaccatgaacgtcccttccaa	420
Db	621	GGGTGGGGGCTTCTACCGGAGCTACACACGAGCTGACCATGACATCCCTTCCAG	680
Qy	421	gccattcacttcattgaactatgaattcctgcaggagcactttaacccccccagagcggtag	480
Query Match	35.2%;	Score 281;	DB 10; Length 5869;
Best Local Similarity	70.1%;	Pred. No. 5.7e-77;	

Db	681	TCCATCCACTTCATCACCTATGAGTTCTGCAGGAGCAGGTCAACCCCGGACCTAC	740
Qy	481	aaccaagctcccgctcctctgagctgagcagagctgagctgagcagcagccaca	540
Db	741	AACCGGAGTCCACATCATCTCAGCGGGCTGGCCGGGGCCCTCGCGCGCGGCCACG	800
Qy	541	accacactggagcttggcaaacactgctcaacacccagggagctcttggc---tttgaac	597
Db	801	ACCCCTGGAGCTCTGTAAGACCTTCTGAACACTCAGGAGAACGTGGCCCTTCGCTG	860
Qy	598	tcacattacagacatatcacaggcatggctagcttagcttcacagagcgtatataagta	657
Db	861	GCCAACTACACGCGCGCTCTCGGTATGCGCAATGCTTCCGACGCTGTACAGTC	920
Qy	658	ggtgggtgacgcctatttccgaggggtgagccagagtaattaccagatcccccc	717
Db	921	AACGGCTG-CGGCTACTTCAAGGCATCCAGGCGCGTGTCTACCATGATGCCCTCC	979
Qy	718	acgccatcgatgctgtgtatgatttcttcaaatcaactaaatacaataaaggca	773
Db	980	ACCGCAATTTCTGCTGTCTATGAGTTCTCAAGTACTTCTCACCACGCGCA	1035
RESULT	15		
LOCUS	AF361699	5869 bp	linear ROD 07-NOV-2001
DEFINITION	Mus musculus mitochondrial carrier-like protein mRNA, complete cds;		
	nuclear gene for mitochondrial product.		
ACCESSION	AF361699		
VERSION	AF361699.1	GI:16755527	
KEYWORDS	house mouse.		
SOURCE	Mus musculus		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE	1 (bases 1 to 5869)		
AUTHORS	Li, Q., Eckenrode, S., Wang, C., Ruan, Q., Shi, J., McIndoe, R.A. and She, J.		
TITLE	A novel mouse mitochondrial carrier protein gene is up-regulated from young to adult NOD mice		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 5869)		
AUTHORS	Li, Q., Eckenrode, S., Wang, C., Ruan, Q., Shi, J., McIndoe, R.A. and She, J.		
TITLE	Direct Submission		
JOURNAL	Submitted (15-MAR-2001) Department of Pathology, Immunology and Laboratory Sciences, University of Florida, 1600 SW Archer Road, Room D6-15, Gainesville, FL 32610, USA		
FEATURES	Location/Qualifiers		
source	1..5869		
	/organism="Mus musculus"		
	/strain="C57BL/6J"		
	/db_xref="taxon:10090"		
	/map="between D14Mit124 and D14Mit236"		
	116..664		
	/note="MCLP"		
	/codon_start=1		
	/evidence=not_experimental		
	/product="mitochondrial carrier-like protein"		
	/protein_id="AAL27990.1"		
	/db_xref="GI:16755528"		
	/translation="MELRGGVGNQAARRMDGSDRGSGGSKDAGSEYENLPTSAS VSTMTAGMAGILEHSIMYPVDSVKTRMQLNDVPHOGNSHLANGIILKAFVWSWEALL PLRLNVMVMGAPAHAMYPACVENMKRTLNDVPHOGNSHLANGIILKAFVWSWEALL SGASSPEPSNLHPRTENSRVT"		
BASE COUNT	1387 a 1427 c 1390 g 1665 t		
ORIGIN			

```
Matches 393; Conservative 0; Mismatches 165; Indels 3; Gaps 1;
Qy 223 ggtgcggccgggtgtgtggcaacattacttcattgatgcagccatgaacctcgggaagtg 282
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2311 GGGGTAGCTGGGAGTATGGGCACCTACTCCACGATCAGTAATGAATCCAGCAGAAAGTG 2370
Qy 283 gtcaagcagaggatgcagatgtacaaactcaccataccaccgggtgacagactgtgtacgg 342
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2371 GTGAACACAGCGCTTACAGATGTACAACTCCAGCACCAGTCAGCCTTCAGTTGTATCCGG 2430
Qy 343 gcagtgtggcaaaatgaaggggccgggcccctttaccgcgagctacaccacccagctgacc 402
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2431 ACAGTGTGGGGACCGAGGGGTGGGGCCCTTCTACAGGAGTTACACACACAGCTGACC 2490
Qy 403 atgaacgttcccttccaaagccattcaacttcattgatgacctatgaattcctgcaggagcacttt 462
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2491 ATGAATATCCCTTTCAGTCAATTCACCTTATCATCACTATGAGTTTCTGCAGGAGCAAGTC 2550
Qy 463 aacccccagagacggtacaaacccaagctcccaagctcctctctgagcttgcgcaggagct 522
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2551 AACCTCGCCGGGACTACAAACCCACAGCTCTCACATCATCTCAGGAGGCTGGCGGGGCA 2610
Qy 523 gtagctgccgcagcccaacacccactggagcttgcacaaacactgctcaacacccagagg 582
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2611 CTGCCGCGAGCTGCCACGACCCCGCTGGAGCTGTGCAAAACCCCTCCTCAACACGCGAGGAG 2670
Qy 583 tccttgctttgaac---tcacacattacaggacacatcacaggcatggttagtgcttc 639
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2671 AACATGGCTCTCTCCCTGGCCACAGTCAGCGCGCGGCTGTGCGGCATGGGCAATGCGCTTC 2730
Qy 640 aggacggtatatcaagtaggtgggggtgacccgctattccgagggggtgcaggccagagta 699
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2731 CGGACGGGTACACAGCTCAACGGGCTTGGCGGCTATTTCAAGGCATCCAGGCTCGAGTC 2790
Qy 700 attaccagatccccctccacagccatcgatgctgtgtatgagttcttcaaataccta 759
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2791 ATCTACCAGATGCCTTCCACCGCATCTCCTGGTCCGCTTTATGAGTTCTTCAAGTACATC 2850
Qy 760 atcaactaaaaggcaagaagag 780
    ||| ||| ||| ||| |||
Db 2851 CTTACAAAGAGGCGAGCTGGAG 2871
```

Search completed: September 28, 2002, 07:57:11
Job time: 14752 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2002, 08:05:59 ; Search time 664.31 Seconds
(without alignments)
2062.437 Million cell updates/sec

Title: US-09-870-113-7
Perfect score: 798
Sequence: 1 atgcagagtctacagcctga.....agtggaggctggcaagtga 798

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_032802.*
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	798	100.0	1294	22	AAI60661 Human polynucleoti
2	798	100.0	1322	22	AAF27733 Human transport pr
3	798	100.0	1336	21	AAC90457 Human uncoupling p
4	790	99.0	1316	22	AAI58875 Human polynucleoti
5	695	87.1	1418	21	AAI74843 Human OREF ORF398
6	639	80.1	710	20	AAZ15876 Human gene express
7	576	72.2	2502	22	AAS03906 Human secreted pro
8	556	69.7	2037	24	AAS62787 cDNA sequence #574
9	520	65.2	997	22	AAH22162 Human mitochondria

10	508	63.7	27960	22	AAK69779 Human immune/haema
11	508	63.7	27960	22	AAK73320 Human immune/haema
12	417.8	52.4	483	24	AB199871 Mouse ischaemic co
13	392	49.1	401	22	AAF66156 Novel human polynu
14	376	47.1	1716	21	AACT77189 Human OREF ORF2744
15	365.6	45.8	1305	22	AAF59920 Human mitochondria
16	300	37.6	300	20	AAZ13026 Human gene express
17	281.2	35.2	1913	22	AAH89944 Human bone marrow
18	277.2	34.7	1814	22	AAK94871 Human full-length
19	276	34.6	1015	22	AAK92501 Human mitochondria
20	276	34.6	1673	22	AAK53156 Human polynucleoti
21	276	34.6	2562	22	AAH90057 Human bone marrow
22	276	34.6	2710	21	AAK66452 Human secreted pro
23	275.6	34.5	1890	22	AAK94897 Human full-length
24	272.2	34.1	600	22	AAH22163 Human mitochondria
25	267.6	33.5	1902	22	AAK52172 Human polynucleoti
26	250	31.3	8211	22	AAK80490 Human immune/haema
27	174	21.8	339	22	AAK58672 Human immune/haema
28	164.2	20.6	792	21	AACT77177 Human OREF ORF2732
29	161.4	20.2	1548	23	AAS65357 DNA encoding novel
30	159.6	20.0	504	23	AAS65354 DNA encoding novel
C 31	153.6	19.2	338	22	AAS23288 Human prostate can
32	141.4	17.7	2137	23	ABL08627 Drosophila melanog
33	138.2	17.3	975	21	AACT77173 Human OREF ORF2728
34	114.8	14.4	3028	22	AAK91320 Human polynucleoti
C 35	113.4	14.2	599	22	AAS24885 Human ovarian PCR-
C 36	113.4	14.2	599	22	AAH83525 Human ovarian tumo
37	105	13.2	842	22	AAK92357 Human cDNA 5'-end
38	105	13.2	842	22	AAK93895 Human cDNA clone r
39	105	13.2	2837	22	AAK80493 Human immune/haema
40	101.8	12.8	1468	21	AACT77175 Human OREF ORF2730
41	101	12.7	622	22	AAK91764 Human cDNA 5'-end
42	101	12.7	622	22	AAK93200 Human cDNA clone r
43	91.4	11.5	1758	20	AAV84573 Human secreted pro
44	91.4	11.5	1758	22	ABA83356 Human secreted pro
45	88	11.0	257	22	AAS31106 Human diagnostic a

ALIGNMENTS

RESULT 1
AAI60661/c
ID AAI60661 standard: cDNA: 1294 BP.
XX AC AAI60661;
XX DT 22-OCT-2001 (first entry)
XX DE Human polynucleotide SEQ ID NO 4650.
XX KW Human: nectropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX OS Homo sapiens.
XX PN WO200153312-A1.
XX PD 26-JUL-2001.
XX PF 26-DEC-2000; 2000WO-US34263.
XX PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.

PR 29-NOV-2000; 2000US-0727344.
PA (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX WPI: 2001-442253/47.
DR P-PSDB: AAM41505.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Claim 1; SEQ ID NO 4650; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 1294 BP; 267 A; 335 C; 368 G; 324 T; 0 other;

Query Match 100.0%; Score 798; DB 22; Length 1294;
Best Local Similarity 100.0%; Pred. No. 2.5e-245;
Matches 798; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgcagagctacagcctgacccagctgcccgcctatgcgaatgctgttgagagccctctgg 60
DB 1134 ATGCAGAGCTACAGCCCTGAGCCAGCTGCCGCTATCCGAATGTTGGAGGCCCTCTGG 1075

QY 61 aggtattataagaagcagggcctatggggccatgagggcctggaacgtcacagcaaca 120
DB 1074 AGGATTATAAGACAGGAGGGCTATGAGAGGCCCATGAGGGGCTGAACCTCACAGCACACA 1015

QY 121 ggcgcagggcctgccacgcctcttatttgcctgctacgaaaagttaaaaaaacattg 180
DB 1014 GCGCAGGGCCTGCCACGCCTTTATTTGCTGCTACGAAAAAGTTAAAAAGACATTG 955

QY 181 agtgaatgaatccaccctgggggcaatagcgcataatggccaatggtggccgggtgtgtg 240
DB 954 AGTGATGTAATCCACCCCTGGGGGCAATAGCCATATGTCGAATGGTGGCGCGGGTGTGTG 895

QY 241 gcaactacttactatgatgcagccataaacctcgaaagtgtcaacgacagatgcag 300
DB 894 GCAACATTACTTCATGATGTCAGCATGAACCTCGGAAGTGGTCAACGAGGATGCAG 835

QY 301 atgtacaactcaccataccaccgggtgacagactgtgtacgggagtggtgcaaaatgaa 360
DB 834 ATGTACAACCTCACCATACCACCGGTGACAGACTGTGTACGGCAGTGTGCAAAATGAA 775

QY 361 gggccggggcctttaccgcagctacacccaccagctgaccatgaacgttccttccaa 420
DB 774 GGGCCGGGGCCTTTTACCAGAGTACACACCACCGAGCTGACCATGATGTCCTTCCAA 715

QY 421 gccattacttcagcctatgaattcctctcagagacacttaacccccagagcgtac 480
DB 714 GCCATTCACTTCATGACCTATGAATTCCTGCAGGAGCACTTTAACCCCCAGAGCGGTAC 655

QY 481 aacccaagctcccacgtctctctctgtgagcttgcgcagagagctgtagctgcccagccaca 540

Db 654 AACCCAGCTCCACGCTCTCTCTGGAGCTTGGCAGGAGCTGTAGCTGCCGAGCCACA 595
QY 541 accccactggcgtttgcaaaaaactgctcaaacccagggagctcttggtttgaactca 600
Db 594 ACCCCACTGGAGCTTGGCAAAACACTGCTCAACACCCAGGAGTCTTGGCTTTGAACCTCA 535
QY 601 cacattacagacatatcacagggcatggctagtgcttccagcagcggtatatacaagtaggt 660
Db 534 CACATTACAGGACATATCACAGGCATGGCTAGTGGCTTCAGGACGGTATATCAAGTAGGT 475
QY 661 ggggtgaccgcctatttccgaggggtgcagggccagagtaattaccagatccccctccaca 720
Db 474 GGGGTGACCGCTATTTCGAGGGGTGCAGGCCAGAGTAATTTACCAGATCCCCCTCCACA 415
QY 721 gccatcgcatgctgtgtatgatgattcttcaaatcacctaaataaagcagaagaag 780
Db 414 GCCATCGCATGCTGTGTATGATGTTCTCAATACCTTAATCACTAAAGGCAAGAAGAG 355
QY 781 tggagggcctgccaagtga 798
Db 354 TGGAGGGCTGCGCAAGTGA 337

RESULT 2
AAF27733
ID AAF27733 standard; cDNA; 1322 BP.
XX
AC AAF27733;
XX
DT 28-MAR-2001 (first entry)
XX
DE Human transport protein TPPT-33 coding sequence.
XX
KW Human; transport protein; TPPT; transport disorder; metabolic disorder;
KW neurological disorder; cardiovascular disorder; reproductive disorder;
KW immune disorder; cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200078953-A2.
XX
PD 28-DEC-2000.
XX
PF 16-JUN-2000; 2000WO-US16668.
XX
PR 17-JUN-1999; 99US-0139923.
PR 10-AUG-1999; 99US-0148177.
PR 18-AUG-1999; 99US-0149357.
PR 28-OCT-1999; 99US-0162287.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Lal P, Yang J, Yue H, Hillman JL, Tang YT, Bandman O, Burford N;
PI Baughn MR, Azimzai Y, Lu DAM, Au-Young J, Patterson C;
XX
XX WPI: 2001-041424/05.
DR P-PSDB: AAB60113.
XX
PT Isolated polypeptide with a human transport protein sequence is useful
PT for the diagnosis, prevention and treatment of disorders associated
PT with the immune, reproductive and cardiovascular systems -
XX
PS Claim 5; Page 160; 165pp; English.
XX
CC The present invention provides the protein and coding sequences for 43
CC novel human transport proteins (designated TPPTs). These can be used in
CC the diagnosis and treatment of transport, metabolic, neurological,
CC reproductive, cardiovascular and immune disorders, and cell proliferative
CC disorders such as cancer.
XX
SQ Sequence 1322 BP; 328 A; 376 C; 351 G; 267 T; 0 other;

```
Query Match      100.0%; Score 798; DB 22; Length 1332;
Best Local Similarity 100.0%; Pred. No. 2.6e-245;
Matches 798; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgcagagctacagcctgaccagctgcccgcctatcgcaatgtgtgagccctctgg 60
   |||||||
Db 186 atgcagagctacagcctgaccagctgcccgcctatcgcaatgtgtgagccctctgg 245

QY 61 aggattataagaacgagggcctatggagcccatgaggggctgaacgtcacacaaca 120
   |||||||
Db 246 aggattataagaacgagggcctatggagcccatgaggggctgaacgtcacacaaca 305

QY 121 ggcgagggcctgcccagcccttattttgctgctacgaaagttaaaaaagacattg 180
   |||||||
Db 306 ggcgagggcctgcccagcccttattttgctgctacgaaagttaaaaaagacattg 365

QY 181 agtgaatgaataccacccctgggggcaatagccatattgccaatgtgcggcggtgtgtg 240
   |||||||
Db 366 agtgaatgaataccacccctgggggcaatagccatattgccaatgtgcggcggtgtgtg 425

QY 241 gcaacattacttcatgatgcagccatgaacctgcggagtggtcaagcagaggtacgag 300
   |||||||
Db 426 gcaacattacttcatgatgcagccatgaacctgcggagtggtcaagcagaggtacgag 485

QY 301 atgtacaactaccataccaccgggtgacagactgtgtacgggagtggtggcaaatgaa 360
   |||||||
Db 486 atgtacaactaccataccaccgggtgacagactgtgtacgggagtggtggcaaatgaa 545

QY 361 ggggcccgggcttttacgcagctacaccacccagctgaccatgaacgttcttccaa 420
   |||||||
Db 546 ggggcccgggcttttacgcagctacaccacccagctgaccatgaacgttcttccaa 605

QY 421 gcattacttcatgacctataatctctgcaggagcactttaacccacagacgggtac 480
   |||||||
Db 606 gcattacttcatgacctataatctctgcaggagcactttaacccacagacgggtac 565

QY 481 aaccaagctccacgtctctctgtgagcttgctgcgagagctgtagctgcgcagccaca 540
   |||||||
Db 666 aaccaagctccacgtctctctgtgagcttgctgcgagagctgtagctgcgcagccaca 725

QY 541 acccactggagcttgcgaaacacactgctcaacacccagagctccttggcttgactca 600
   |||||||
Db 726 acccactggagcttgcgaaacacactgctcaacacccagagctccttggcttgactca 785

QY 601 cacattacaggacatatcacaggcagtgctagtgcttcagagcaggtatatcaagtggt 660
   |||||||
Db 786 cacattacaggacatatcacaggcagtgctagtgcttcagagcaggtatatcaagtggt 845

QY 661 ggggtgacgcctatttcgaggggtgagggccagagtaatttacagatccctccaca 720
   |||||||
Db 846 ggggtgacgcctatttcgaggggtgagggccagagtaatttacagatccctccaca 905

QY 721 gccatcgcatggtctgtgtatgatttctcaaatcccttaatacctaaaggcaagaagag 780
   |||||||
Db 906 gccatcgcatggtctgtgtatgatttctcaaatcccttaatacctaaaggcaagaagag 965

QY 781 tggagggtgccaagtga 798
   |||||||
Db 966 tggagggtgccaagtga 983

RESULT 3
ID AAC90457
XX AAC90457 standard; cDNA; 1336 BP.
AC AAC90457;
XX 12-MAR-2001 (first entry)
XX Human uncoupling protein CDNA #6.
XX Human; uncoupling protein; immunosuppressive; antiarthritic;
```

Qy 361 ggggcccgggctttaccgcagctacaccacccagctgaccatgaacgttccttcccaa 420
Db 557 ggggcccgggctttaccgcagctacaccacccagctgaccatgaacgttccttcccaa 616
Qy 421 gccattcactcatgacctatgaattcctgcaggagcactttaaccccccagagcgggtac 480
Db 617 gccattcactcatgacctatgaattcctgcaggagcactttaaccccccagagcgggtac 676
Qy 481 aacccaaagctccacgctctctgagcttgcgcagagcgtgtagctgcccgcagccaca 540
Db 677 aacccaaagctccacgctctctgagcttgcgcagagcgtgtagctgcccgcagccaca 736
Qy 541 accccactggagctttgcaaacactgctcaacacccagggagtccttggctttgaaactca 600
Db 737 accccactggagctttgcaaacactgctcaacacccagggagtccttggctttgaaactca 796
Qy 601 cacattacaggacatatcacaggcatggtctagtgcccttcaggagcgggtatatacagtaggt 660
Db 797 cacattacaggacatatcacaggcatggtctagtgcccttcaggagcgggtatatacagtaggt 856
Qy 661 ggggtgaccgctctattccgaggggtgcaggccagagtaattaccagatccctccaca 720
Db 857 ggggtgaccgctctattccgaggggtgcaggccagagtaattaccagatccctccaca 916
Qy 721 gccatcgatggtctgtgtatgagttcttcaaatcacctaatcaactaaaggcaagagag 780
Db 917 gccatcgatggtctgtgtatgagttcttcaaatcacctaatcaactaaaggcaagagag 976
Qy 781 tggaggggctggcaagtga 798
Db 977 tggaggggctggcaagtga 994

RESULT 4

AAI58875
ID AAI58875 standard; cDNA; 1316 BP.
XX AC AAI58875;
XX DT 22-OCT-2001 (first entry)
XX DE Human polynucleotide SEQ ID NO 1078.
XX KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX OS Homo sapiens.
XX PN WO200153312-A1.
XX PD 26-JUL-2001.
XX PF 26-DEC-2000; 2000WO-US34263.
XX PR 21-JAN-2000; 2000US-0488725.
XX PR 25-APR-2000; 2000US-0552317.
XX PR 09-JUL-2000; 2000US-0598042.
XX PR 19-JUL-2000; 2000US-0620312.
XX PR 03-AUG-2000; 2000US-0653450.
XX PR 14-SEP-2000; 2000US-0662191.
XX PR 19-OCT-2000; 2000US-0693036.
XX PR 29-NOV-2000; 2000US-0727344.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

DR WPI; 2001-442253/47.
DR P-PSDB; AAM39719.
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX Claim 1; SEQ ID NO 1078; 10078pp; English.
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC C.N.S disorders
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX SQ Sequence 1316 BP; 335 A; 364 C; 337 G; 280 T; 0 other;
Query Match 99.08; Score 790; DB 22; Length 1316;
Best Local Similarity 99.4%; Pred. No. 9.4e-243;
Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1 atgcagagctcacagcctgaccagctgcccgcctatgcgaatgtgttgaggccctctgg 60
Db 176 atgcagagctcacagcctgaccagcctgcccgcctatgcgaatgtgttgaggccctctgg 235
Qy 61 aggtattataagaacggaggccctatgaggcccatgagggggctgaacgtcacacagcaaca 120
Db 236 aggtattataagaacggaggccctatgaggcccatgagggggctgaacgtcacacagcaaca 295
Qy 121 ggcgaggggcctgccacgcctttatttgcctgctacgaaagttaaaaaagacattg 180
Db 296 ggcgagctgctgccacgcctttatttgcctgctacgaaagttaaaaaagacattg 355
Qy 181 agtcatgtaatccacctgggggcaatagccatattgccaatgtgcccggcgtgtgtg 240
Db 356 agtcatgtaatccacctgggggcaatagccatattgccaatgtgcccggcgtgtgtg 415
Qy 241 gcaacattacttcattgatgcagccatgaacccctgcgaagtgggtcaagcagagatgcag 300
Db 416 gcaacattacttcattgatgcagccatgaacccctgcgaagtgggtcaagcagagatgcag 475
Qy 301 atgtacaactcacatcacccgggtgacagactgtgtacgggagctgtggcacaatgaa 360
Db 476 atgtacaactcacatcacccgggtgacagactgtgtacgggagctgtggcacaatgaa 535
Qy 361 ggggcccgggctttaccgcagctacaccacccagctgaccatgaacgttccttcccaa 420
Db 536 ggggcccgggctttaccgcagctacaccacccagctgaccatgaacgttccttcccaa 595
Qy 421 gccattcactcatgacctatgaattcctgcaggagcactttaaccccccagagcgggtac 480
Db 596 gccattcactcatgacctatgaattcctgcaggagcactttaaccccccagagcgggtac 655
Qy 481 aacccaaagctccacgctctctctctgcaggcttgcgcagagcgtgtagctgcccgcagcaaca 540
Db 656 aacccaaagctccacgctctctctctgcaggcttgcgcagagcgtgtagctgcccgcagcaaca 715
Qy 541 accccactggagctttgcaaacactgctcaacacccagggagtccttggctttgaaactca 600
Db 716 accccactggagctttgcaaacactgctcaacacccagggagtccttggctttgaaactca 775
Qy 601 cacattacaggacatatcacaggcatggtctagtgcccttcaggagcgggtatatacagtaggt 660

Db 776 cacattacaggacatcacaggcatggtgtagtgccttcagagcggtatatacaagtagt 835
 Qy 661 ggggtgaccgcctatttcagagggtgagccagagtaattaccagatccctccaca 720
 Db 836 ggggtgaccgcctatttcagagggtgagccagagtaattaccagatccctccaca 895
 Qy 721 gccatgcgatgctgtgatgatttcttcaaatcacctaatcaactaaaggcaagaag 780
 Db 896 gccatgcgatgctgtgatgatttcttcaaatcacctaatcaactaaaggcaagaag 955
 Qy 781 tggagggtgcaagtga 798
 Db 956 tggagggtgcaagtga 973
 RESULT 5
 AAC74843
 ID AAC74843 standard; cDNA; 1418 BP.
 XX AC AAC74843;
 XX DT 08-FEB-2001 (first entry)
 XX DE Human ORFX ORF398 polynucleotide sequence SEQ ID NO:795.
 XX KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnery; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; vasotrophic; antidiabetic;
 KW immunostimulant; thrombolytic; coagulant; immunosuppressant; cardiant;
 KW hypotensive; dermatological; immunosuppressive; antidiabetic;
 KW antiviral; antibacterial; antifungal; antirheumatic; antihypertoid;
 KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KW thrombosis; contraceptive; ss.
 XX OS Homo sapiens.
 XX PN WO200058473-A2.
 XX PD 05-OCT-2000.
 XX PF 31-MAR-2000; 2000WO-US08621.
 XX PR 31-MAR-1999; 99US-0127607.
 XX PR 02-APR-1999; 99US-0127636.
 XX PR 05-APR-1999; 99US-0127728.
 XX PR 30-MAR-2000; 2000US-0540763.
 XX PA (CURA-) CURAGEN CORP.
 XX PI Shimkets RA, Leach M;
 XX WPI: 2000-602362/57.
 XX DR P-PSDB; AAB40634.
 XX PT Novel nucleic acids and peptides derived from open reading frame x,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX Claim 5; Page 853-854; 5507pp; English.
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
 CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotrophic;

CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
 CC antihypertoid; and antianaemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.
 XX Sequence 1418 BP; 353 A; 400 C; 368 G; 297 T; 0 other;

Query Match 87.1%; Score 695; DB 21; Length 1418;
 Best Local Similarity 89.6%; Pred. No. 2.9e-212;
 Matches 798; Conservative 0; Mismatches 0; Indels 93; Gaps 1;

Qy 1 atgcagagtctacagcctgacccagctgcgcgctatcgcaatgtgttgaggccctctgg 60
 Db 178 atgcagagtctacagcctgacccagctgcgcgctatcgcaatgtgttgaggccctctgg 237
 Qy 61 aggattataaagcagaggccctatggagcccatgagggcgctgaacgtcacagcaaca 120
 Db 238 aggattataaagcagaggccctatggagcccatgagggcgctgaacgtcacagcaaca 297
 Qy 121 ggcgaggcctgccacgcctttatttgcctgctacgaaagttaaaaaagacattg 180
 Db 298 ggcgaggcctgccacgcctttatttgcctgctacgaaagttaaaaaagacattg 357
 Qy 181 agtgatgataccacctgggggcaatagccatattgccaatggtgcggccgggtgtgtg 240
 Db 358 agtgatgataccacctgggggcaatagccatattgccaatggtgcggccgggtgtgtg 417
 Qy 241 gcaacattacttcatgctgacgcctgaacccctgcgga----- 278
 Db 418 gcaacattacttcatgctgacgcctgaacccctgcggaagcgtgctgctgactgggg 477
 Qy 279 ----- 278
 Db 478 ctctgaatctgatactctccatcacccggttggtgctgtccaccattctctccctgtg 537
 Qy 279 -----agtgtcaagcagaggtgagatgtacaaactacacacacgggtg 327
 Db 538 atggcactactagtgtcaagcagaggtgagatgtacaaactacacacacgggtg 597
 Qy 328 acagactgttacgggcagtggtgcacaaatgaaggcgccgggctttttaccgcagctac 387
 Db 598 acagactgtgtacgggcagtggtgcacaaatgaaggcgccgggctttttaccgcagctac 657
 Qy 388 accaccagctgacatgaacgttctcttccaagccattcacttcatgacctatgaattc 447
 Db 658 accaccagctgacatgaacgttcttccaagccattcacttcatgacctatgaattc 717
 Qy 448 ctgcaggagcatttaaccccccagagacggtacaaacccaagctccacgctctctctgga 507
 Db 718 ctgcaggagcatttaaccccccagagacggtacaaacccaagctccacgctctctctgga 777
 Qy 508 gcttgcgcagagctgtagctgcgcagcacaacccccacacggagcttttgaacacactg 567
 Db 778 gcttgcgcagagctgtagctgcgcagcacaacccccacacggagcttttgaacacactg 837
 Qy 568 ctcaacacccaggagctctcttggtctttgaactcacacattacaggagcatatcacaggcatg 627
 Db 838 ctcaacacccaggagctctcttggtctttgaactcacacattacaggagcatatcacaggcatg 897
 Qy 628 gctagtgcttccagagcgttatatacaagtagtggtgggtgaccgcctatttccaggggtg 687

Db 898 gtagtgcttcaggacggtatatcaagtagtgggtgaccgctatttccgaggggtg 957
Qy 688 caggccagagtaattaccagatccctccacagccatcgcatggtctgtgtatgagttc 747
Dd 958 caggccagagtaattaccagatccctccacagccatcgcatggtctgtgtatgagttc 1017
Qy 748 ttcaaatccttaatacctaaaggaagagtagtggaggtggtggaagtga 798
Dd 1018 ttcaaatccttaatacctaaaggaagagtagtggaggtggtggaagtga 1068
RESULT 6
AAZ15876
ID AAZ15876 standard; cDNA; 710 BP.
XX
AC AAZ15876;
XX
DT 12-OCT-1999 (first entry)
XX
DE Human gene expression product cDNA sequence SEQ ID NO:3345.
XX
KW Human; gene: gene expression product; diagnosis; therapy; probe;
KW detection; mapping; tissue typing; profiling; forensic; cancer;
KW genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.
XX
OS Homo sapiens.
XX
PN W09938972-A2.
XX
PD 05-AUG-1999.
XX
PF 28-JAN-1999; 99WO-US01619.
XX
PR 03-APR-1998; 98US-0080666.
PR 28-JAN-1998; 98US-0072910.
PR 24-FEB-1998; 98US-0075954.
PR 31-MAR-1998; 98US-0080114.
PR 03-APR-1998; 98US-0080515.
XX
PA (CHIR) CHIRON CORP.
PA (HYSE) HYSEQ INC.
XX
PI Crkvenjakov R, Dickson M, Drmanac R, Drmanac S;
PI Escobedo J, Garcia PD, Garcia V, Giese K, Innis MA;
PI Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;
PI Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;
PI Stache-Crain B, Sudduth-Klinger J, Williams LT;
XX
DR WPI; 1999-494092/41.
XX
PT Novel human genes and their expression products which are
PT differentially expressed in different cell types
XX
PS Claim 1; Page 1602; 2479pp; English.
XX

The present invention describes a library of human polynucleotides comprising the sequences given in AAZ12532 to AAZ1779. Also described is a method of detecting differentially expressed genes correlated with the cancerous state of a mammalian cell, comprising detecting at least one differentially expressed gene product in a test sample from a cell suspected of being cancerous, where the gene product is encoded by one of the 5248 polynucleotide sequences given in AAZ12532 to AAZ1779. The polynucleotides can be used as a source of primers and probes, which can be used for a variety of purpose, e.g. detection of expression levels, mapping, tissue typing or profiling, forensics, genetic analysis and detection of polymorphisms. Polypeptides encoded by the polynucleotides can be used for raising antibodies for experimental, diagnostic and therapeutic purposes. The polynucleotides may also be used to construct arrays for diagnostics (which may be used to determine function of an encoded protein); and to detect differences in expression levels between two cells (e.g. to identify abnormal or diseased tissue in a human, to identify a genetic predisposition or susceptibility to a disease such as cancer). The polynucleotides of the invention are especially used in the

CC diagnosis, prognosis and management of colorectal cancer, breast cancer,
CC and lung cancer. The polynucleotides can also be used to screen for
CC peptide analogues and antagonists.
XX
SQ Sequence 710 BP; 186 A; 187 C; 183 G; 154 T; 0 other;
Query Match 80.1%; Score 639; DB 20; Length 710;
Best Local Similarity 100.0%; Pred. No. 1.9e-194;
Matches 639; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 160 gaaaagttaaaaaagacattgagtgatgaatccacccctgggggcaatagccattgtgcc 219
Dd 46 gaaaagttaaaaaagacattgagtgatgaatccacccctgggggcaatagccattgtgcc 105
Qy 220 aatggtgcggccgggtgtgtggcaacatttacttactatgacagccatgaacccgcgaaa 279
Dd 106 aatggtgcggccgggtgtgtggcaacatttacttactatgacagccatgaacccgcgaaa 165
Qy 280 gtggtcaagcagagagatgcagatgtacaactccacataccacccgggtgacagactgtgta 339
Dd 166 gtggtcaagcagagagatgcagatgtacaactccacataccacccgggtgacagactgtgta 225
Qy 340 cgggcagtggtggcaaaatgaaggccggggcctttaccgcagctacacccaccagctg 399
Dd 226 cgggcagtggtggcaaaatgaaggccggggcctttaccgcagctacacccaccagctg 285
Qy 400 accatgaacgttcttccaaagccatttccatgacccatgacccatgacccatgacccatg 459
Dd 286 accatgaacgttcttccaaagccatttccatgacccatgacccatgacccatgacccatg 345
Qy 460 ttaacccccagagacggtacacaccccaagctccacagctccctctctgagcttgcgcagga 519
Dd 346 ttaacccccagagacggtacacaccccaagctccacagctccctctctgagcttgcgcagga 405
Qy 520 gctgtagctgcgcagccacacaccccccactgacgcttttgcaaaacactgctcaacacccag 579
Dd 406 gctgtagctgcgcagccacacaccccccactgacgcttttgcaaaacactgctcaacacccag 465
Qy 580 gagtccctggcttgaactcaacattacagacatatccacagcattgagctagtgcttc 639
Dd 466 gagtccctggcttgaactcaacattacagacatatccacagcattgagctagtgcttc 525
Qy 640 aggcaggtatatacaagtagtggtgggtgacccctatttccagaggtgacagccagagta 699
Dd 526 aggcaggtatatacaagtagtggtgggtgacccctatttccagaggtgacagccagagta 585
Qy 700 attaccagatccccctccacagccatgcgcatggtgtgtatgagttcttcaataactta 759
Dd 586 attaccagatccccctccacagccatgcgcatggtgtgtatgagttcttcaataactta 645
Qy 760 atcaactaaaaggcaagaagtagtgagggctggcgaagtga 798
Dd 646 atcaactaaaaggcaagaagtagtgagggctggcgaagtga 684

RESULT 7
AAS03906
ID AAS03906 standard; cDNA; 2502 BP.
XX
AC AAS03906;
XX
DT 29-AUG-2001 (first entry)
XX
DE Human secreted protein gene #25.
XX
KW Human secreted protein; autoimmune disorder; hyperproliferative disorder;
KW cardiovascular disorder; cerebrovascular disorder; angiogenesis;
KW nervous system disorder; bacterial infection; viral infection; ss;
KW fungal infection; ocular disorder; wound healing; tissue regeneration;
KW epithelial cell proliferation; skin ageing; chemotaxis; I9G fc region.
XX
OS Homo sapiens.

XX WO200123598-A1.
PN 05-APR-2001.
PD 26-SEP-2000; 2000WO-US26324.
PF 27-SEP-1999; 990US-0155807.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Komatsoulis G, Ruben SM, Rosen CA;
PI WPI; 2001-281684/29.
DR P-PSDB; AAU01950, AAU01989.
XX Forty one nucleic acid molecules encoding human secreted proteins, useful
PT in the prevention, treatment and diagnosis of cancer, immune disorders,
PT cardiovascular disorders and neurological diseases -
XX
XX Disclosure: Page 460-461; 518pp; English.
XX Sequences AAS03873-AAS03922 represent isolated nucleic acid molecules and
CC PCR primers of the invention. acid of the invention. Secreted proteins
CC and their related nucleic acids can be used in the diagnosis of or
CC susceptibility to a pathological condition by determining the presence or
CC absence of a mutation in a nucleic acid or the presence or amount of
CC expression of a secreted protein. The sequences are used to prevent,
CC treat or ameliorate a medical condition in e.g. humans, mice, rabbits,
CC goats, horses, cats, dogs, chickens or sheep. The antibodies to the
CC polypeptides can also be used in alleviating symptoms associated with
CC disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme
CC linked immunosorbent assays (ELISA). The disorders include autoimmune
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,
CC nervous system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection.
CC The peptides can also be used to aid wound healing and epithelial cell
CC proliferation, to help prevent skin ageing due to sunburn, to maintain
CC organs before transplantation, to regenerate tissues, in chemotaxis and
CC as a food additive or preservative to alter storage capabilities.
XX
XX Sequence 2502 BP; 587 A; 655 C; 669 G; 589 T; 2 other;
SQ

Query Match 72.2%; Score 576; DB 22; Length 2502;
Best Local Similarity 100.0%; Pred. No. 5.7e-174;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 223 ggtcgccgggtgtgtggaacattacttcatgtatgacgcatgaacccctgcggaagt 282
Db 1513 ggtcgccgggtgtgtggaacattacttcatgtatgacgcatgaacccctgcggaagt 1572

QY 283 gtcaagcagagatgcagatgtacaactcaacaccacccgggtgacagactgtgtacgg 342
Db 1573 gtcaagcagagatgcagatgtacaactcaacaccacccgggtgacagactgtgtacgg 1632

QY 343 gcagtggtgcaaaatgaaggccggccgttttaccagcagctacaccaccagctgacc 402
Db 1633 gcagtggtgcaaaatgaaggccggccgttttaccagcagctacaccaccagctgacc 1692

QY 403 atgaacgttcttccaaagccattcaactcatgacctatgaattctcgtcagagcacatt 462
Db 1693 atgaacgttcttccaaagccattcaactcatgacctatgaattctcgtcagagcacatt 1752

QY 463 aacccccagagcgtacaacccaagctccacgtctctctgtggagcttgccagagact 522
Db 1753 aacccccagagcgtacaacccaagctccacgtctctctgtggagcttgccagagact 1812

QY 523 gtagtgcgcagcagcaaacccccactgacgttttgcaaaaactcgtcacaacccagag 582
Db 1813 gtagtgcgcagcagcaaacccccactgacgttttgcaaaaactcgtcacaacccagag 1872

QY 583 tcttggtttgaactcacacattacagacatatcacagcatggttagtgccttcagg 642
Db 1873 tcttggtttgaactcacacattacagacatatcacagcatggttagtgccttcagg 1932

QY 643 acggtatatcaagtagtggtgacccgctatttccaggggtgcagggccagagtaatt 702
Db 1933 acggtatatcaagtagtggtgacccgctatttccaggggtgcagggccagagtaatt 1992

QY 703 taccagatccctccacagcgcacatgcgtgtgtgtatgatttcttcaaatcctaatt 762
Db 1993 taccagatccctccacagcgcacatgcgtgtgtgtatgatttcttcaaatcctaatt 2052

QY 763 actaaaggcaagaagagtgagggtggtggaagtga 798
Db 2053 actaaaggcaagaagagtgagggtggtggaagtga 2088

RESULT 8
AAS62787
ID AAS62787 standard; cDNA; 2037 BP.
XX
AC AAS62787;
XX
DT 14-FEB-2002 (first entry)
XX
DE cDNA sequence #574 encoding novel human secreted protein.
XX
KW Human secreted protein; hyperproliferative disorder; autoimmune disorder;
KW immune deficiency disorder; blood disorder; inflammatory disorder;
KW infectious disorder; gene therapy; antimicrobial; hepatotropic;
KW immunosuppressive; antirheumatic; ss.
XX
OS Homo sapiens.
XX
PN WO200177291-A2.
XX
PD 18-OCT-2001.
XX
PF 29-MAR-2001; 2001WO-US10485.
XX
PR 06-APR-2000; 2000US-195604P.
XX
PA (GEMY) GENETICS INST INC.
XX
PI Wong GG, Clark HF, Fechtel K, Agostino MJ, Howes SH, Resnick RJ;
PI Gulukota K, Graham JR;
XX
DR WPI; 2002-010900/01.
XX
PT New polynucleotides encoding secreted proteins useful for treating e.g.
PT asthma, HIV and Crohn's disease -
XX
PS Claim 1; Page 364-365; 391pp; English.
XX
CC The present invention relates to the isolation of novel cDNA sequences
CC which encode human secreted proteins. The cDNA sequences have been
CC derived from a variety of human tissues. The invention also provides
CC a method for producing proteins from these polynucleotide sequences.
CC The proteins are useful for identifying compounds that modulate their
CC activity and production, and the cell is also useful for identifying
CC compounds that modulate expression of the polynucleotide sequences
CC encoding the secreted proteins. The sequences of the invention are
CC useful for treating diseases such as hyperproliferative disorders
CC (e.g. cancer), immune deficiency disorders (e.g. severe combined
CC immunodeficiency (SCID)), autoimmune disorders (e.g. multiple
CC sclerosis), blood disorders (e.g. thrombocytopaenia), inflammatory
CC disorders (e.g. arthritis) and infectious disorders (e.g. hepatitis).
CC The polynucleotide sequences of the invention are also useful in gene
CC therapy. AAS62214-AAS62838 represent the cDNA sequences of the
CC invention that encode for novel human secreted proteins.
XX
SQ Sequence 2037 BP; 475 A; 522 C; 524 G; 515 T; 1 other;

[illegible]

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX Homo sapiens.
XX WO200157182-A2.
XX 09-AUG-2001.
XX 17-JAN-2001; 2001WO-US01354.
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0232403.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 21-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.


```
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232081.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0239035.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-024617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.

PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and
XX metastasis -
XX
XX Disclosure; SEQ ID NO 28132; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
XX amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patients own production of (I). Additionally, (I)
XX polynucleotides may be used to produce the secreted (I), by inserting
XX the nucleic acids into a host cell and culturing the cell to express the
XX protein. (I) proteins and polynucleotides may be used to prevent,
XX diagnose and treat immune/hematopoietic-related diseases, especially
XX cancers and cancer metastases of haematopoietic-derived cells. AAK64703
XX to AAK87694 represent human immune/hematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK54950 and AAK82169
XX represent sequences used in the exemplification of the present invention.
XX
XX Sequence 27960 BP; 7998 A; 5835 C; 6583 G; 7544 T; 0 other;

Query Match 63.7%; Score 508; DB 22; Length 27960;
Best Local Similarity 99.8%; Pred. No. 1.3e-151;
Matches 519; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 279 agtgggtcaagcagaggtgcagatgtacaactaccatccaccacgggtgcagactgtgt 338
Db 20845 agtgggtcaagcagaggtgcagatgtacaactaccatccaccacgggtgcagactgtgt 20904
Qy 339 acgggcagtggtgcaaaatgaagggcgggggccctttaccgcagctacacacccagct 398
Db 20905 acgggcagtggtgcaaaatgaagggcgggggccctttaccgcagctacacacccagct 20963
Qy 399 gaccatgaacgttctcttccaaagccattcacttcacgttatgaattcctgcaggagca 458
Db 20964 gaccatgaacgttctcttccaaagccattcacttcacgttatgaattcctgcaggagca 21023
Qy 459 ctttaacccccagagcgtgtacaacccccaaagctccacgtctctctgtgagcttgcagg 518
Db 21024 ctttaacccccagagcgtgtacaacccccaaagctccacgtctctctgtgagcttgcagg 21083
Qy 519 agctgtagctccgcagccacacccccactggagcgttttgaaaaaacactgtctcaacacca 578
Db 21084 agctgtagctccgcagccacacccccactggagcgttttgaaaaaacactgtctcaacacca 21143
```

QY	579	ggagtccttggtcttgaaactcaacacattacaaggacatatcacaggcatggtcgtagtgcctt	638
Db	21144	ggagtccttggtcttgaaactcaacacattacaaggacatatcacaggcatggtcgtagtgcctt	21203
QY	639	caggacggtatatacaagtagggtgggagcgcctattccgagggtgcaggccagaagt	698
Db	21204	caggacggtatatacaagtagggtgggagcgcctattccgagggtgcaggccagaagt	21263
QY	699	aattaccagatccctccacagccatcgatggtctgtgtatgagttcttcaaatacct	758
Db	21264	aattaccagatccctccacagccatcgatggtctgtgtatgagttcttcaaatacct	21323
QY	759	aatactaagaagcgaagaagatggagggtggcgaagtga	798
Db	21324	aatactaagaagcgaagaagatggagggtggcgaagtga	21363

RESULT 12

ABI99871

ID ABI99871 standard; cDNA; 483 BP.

AC ABI99871;

DT 07-MAR-2002 (first entry)

DE Mouse ischaemic condition related cDNA sequence SEQ ID NO:1009.

Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;

KW vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.

OS Mus musculus.

PN WO200188188-A2.

22-NOV-2001.

18-MAY-2001; 2001WO-JP04192.

PR 18-MAY-2000; 2000JP-0145977.

PA (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.

PI
xx
Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y:

DR WPI: 2002-034733/04.

Examining the ischemic condition (e.g. occlusive ischemia) by measuring expression levels of particular genes defined in the specification or by determining the expression profile of a gene group comprising these genes -

PS Claim 2: page 2555-2556: 2690pp: English.

The present invention describes a method for examining ischaemic conditions, comprising measuring the expression levels of particular genes (I) in a test sample or determining the expression profile of a gene group in the sample comprising genes selected from (I). The method is useful for examining the ischaemic condition (e.g. compressive ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring expression levels of particular genes (ABI99202 to ABI99912, encoding the protein sequences in AB57020 to AB57374) or by determining the expression profile of a gene group comprising these genes. The expression levels or expression profiles produced by these genes are used as an indicator when screening for ischaemic condition-improving drugs or therapeutics for ischaemic diseases. ABI99913 and ABI99914 represent PCR primers for a mouse ischaemic condition related sequence, which are used in the exemplification of the present invention.

Sequence 483 BP; 127 A; 139 C; 121 G; 96 T; 0 other;

Query Match

52.48: score 417.8: DB 24: Length 483:

Best Local Similarity 93.28; Pred. No. 1.4e-123;

.....

	Matches	437; Conservative	0; Mismatches	32; Indels	0; Gaps	0;
Qy	165	gttaaaaaagacattgagtgatgtaatccaccctgggggcaataggccatatgtgccaatgg	224			
Db	15	gttaaaaaagacattgagtgaactaaaccacccaggggccaataggccatatgtgccaatgg	74			
Qy	225	tggcgccgggtgtgtgcaaacattacttcatatgtgcagcgcattgaacctctgcgaagtcgtt	284			
Db	75	tgcagccggatgtgtgcgacattacttcatatgtgcagcgcattgaatccacggaagtcgt	134			
Qy	285	caagcacagagatgcagatgtacaactcacatacccacccggtagcacagactgtgtacgggc	344			
Db	135	caagcacagagatgcagatgtacaactgcgcgttacccacccgctgacagactgtgttogggc	194			
Qy	345	agtgtggcaaaaagaaggccggcgccctttaccgcagctacacacccagctgaccat	404			
Db	195	agtgtggcaaaaagaaggccggcgccctttaccgcagctacacacccagctgaccat	254			
Qy	405	gaacgttcttccaaagccattcacttcatgaactatgaactatgaactatgaactatgaacttaa	464			
Db	255	gaatgtccccctccaaagccattcacttcatgaactatgaactatgaactatgaactatgaacttaa	314			
Qy	465	ccccacagagcgtgtacaaccccaagctcccaagctctctcttgagcttgcgcaggagctgt	524			
Db	315	ccccacagagcgtgtacaaccccaagctcccaagctctctcttgagcttgcgcaggagcagt	374			
Qy	525	agctgcgcgcagccacaccccaactggagctttgcaaaacactgctcaacacccaggagtc	584			
Db	375	agctgcgcgcgcacaccccaactggagctttgaaacacactgctcaacacccagggaatc	434			
Qy	585	cttggttgtgaatcacacatcacaggacatataccagggcattggctagt	633			
Db	435	cctgtgcttgaactcaaacattactggaacacatcacaggcattggctagt	483			
RESULT	13					
AAF66156	ID	AAF66156 standard; cDNA; 401 BP.				
XX	AC	AAF66156;				
XX	DT	09-APR-2001 (first entry)				
XX	DE	Novel human polynucleotide, SEQ ID NO: 1912.				
XX	KW	Human; cytostatic; gene therapy: colon cancer; prostate cancer;				
XX	KW	breast cancer; lung cancer; cancer detection; ss.				
XX	OS	Homo sapiens.				
XX	PX	WO200102568-A2.				
XX	PD	11-JAN-2001.				
XX	PF	30-JUN-2000; 2000WO-US18374.				
XX	PR	02-JUL-1999; 99US-0142310.				
XX	PR	02-JUL-1999; 99US-0142311.				
XX	PA	(CHIR) CHIRON CORP.				
XX	PA	(HYSE-) HYSEQ INC.				
XX	PI	Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A;				
XX	PI	Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G, Drmanac R;				
XX	PI	Crkenjakov R, Drmanac S, Dickinson M, Labat I, Leshkowitz D;				
XX	DR	Kita D, Garcia V, Jones LW, Strache-Crain B;				
XX	WI	WPI; 2001-091805/10.				
XX	PT	Library of polynucleotides for diagnosing a cancerous state of a				
XX	PT	mammalian cell and detecting cancer, particularly of the colon or				
XX	PT	prostate, comprises 3351 human polynucleotide sequences -				

PS Claim 9; Page 818; 1046pp; English.

XX The present sequence is one of 3351 sequences in a library of human
CC polynucleotides. The library is used to detect differentially expressed
CC genes correlated with a cancerous state of a mammalian cell and can
CC detect colon, prostate, breast and lung cancer. The library can be used
CC to produce probes for detection of mRNA and to produce additional copies
CC of the polynucleotides. The probes can be used for chromosome mapping of
CC the polynucleotide and for detection of transcription levels. Ribozymes
CC or antisense oligonucleotides can be generated. The polynucleotides and
CC their gene products are used as genetic or biochemical markers (e.g. in
CC blood or tissues) that will detect the earliest changes along the
CC carcinogenesis pathway and/or monitor the efficacy of therapies and
CC preventive interventions. The polynucleotides, polypeptides and
CC antibodies against them can be used in pharmaceutical compositions to
CC treat the cancers and proliferative disorders such as neoplasia,
CC dysplasia and hyperplasia.

XX Sequence 401 BP; 100 A; 104 C; 120 G; 76 T; 1 other;

Query Match 49.1%; Score 392; DB 22; Length 401;

Best Local Similarity 100.0%; Pred. No. 2.3e-115;

Matches 392; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 gtctacagcctgacccagctgccgctatcgcaatgtgttgagccctctggaggatta 67

Dy ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 9 gtctacagcctgacccagctgccgctatcgcaatgtgttgagccctctggaggatta 68

Qy 68 taagaacaggaggcctatggaggcccatgaggggctgaacgtcacacagcgcgag 127

Dy ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 69 taagaacaggaggcctatggaggcccatgaggggctgaacgtcacacagcgcgag 128

Qy 128 ggctgccagcccttatttgcctgtactcgaagaatttaaaagacattgagtgtg 187

Dy ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 129 ggctgccagcccttatttgcctgtactcgaagaatttaaaagacattgagtgtg 188

Qy 188 taatcacccctgggggcaatagccatattgccaattgtgcgcgggtgtgtggcaacat 247

Dy ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 189 taatcacccctgggggcaatagccatattgccaattgtgcgcgggtgtgtggcaacat 248

Qy 248 tacttcagtatgcagcctgaacctgcggaagtgtcaacagaggtgcagatgtaca 307

Dy ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 249 tacttcagtatgcagcctgaacctgcggaagtgtcaacagaggtgcagatgtaca 308

Qy 308 actcaccatcacccgggtgcagactgtacggcagctgtgcgaataatgaaggggcgg 367

Dy ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 309 actcaccatcacccgggtgcagactgtacggcagctgtgcgaataatgaaggggcgg 368

Qy 368 gggccttttacgcagctacaccccccagctg 399

Dy ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 369 gggccttttacgcagctacaccccccagctg 400

RESULT 14

AAC77189

ID AAC77189 standard; cDNA; 1716 BP.

XX AAC77189;

AC AAC77189;

XX 08-FEB-2001 (first entry)

DT 08-FEB-2001 (first entry)

XX Human ORFX ORF2744 polynucleotide sequence SEQ ID NO:5487.

DE Human;

XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;

XX vulnary; antipariatic; antiparkinsonian; nootropic; neuroprotective;

KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiac;

KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;

KW hypotensive; dermatological; immunosuppressive; antiinflammatory;

KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;

KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;

KW neurodegenerative disorder; osteoarthritis; graft vs host disease;

KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;

KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
XX thrombosis; contraceptive; ss.

OS Homo sapiens.

XX WO200058473-A2.

PN 05-OCT-2000.

PD 31-MAR-2000; 2000WO-US08621.

PF 31-MAR-1999; 99US-0127607.

PR 02-APR-1999; 99US-0127636.

PR 05-APR-1999; 99US-0127728.

PR 30-MAR-2000; 2000US-0540763.

XX (CURA-) CURAGEN CORP.

XX Shimkets RA, Leach M;

XX WPI; 2000-602362/57.

DR P-PSDB; AAB42980.

XX Novel nucleic acids and peptides derived from open reading frame X,
XX useful for treating e.g. cancers, proliferative disorders,
XX neurodegenerative disorders and cardiovascular disease -

PS Claim 5; Page 4661-4662; 5507pp; English.

XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
XX which represent the human ORFX open reading frames 1 to 3161. The ORFX
XX sequences have activities such as: cytostatic; hepatotropic; vulnary;
XX antipariatic; antiparkinsonian; nootropic; immunoprotective;
XX osteopathic; anticonvulsant; antiarthritic; neuroprotective;
XX immunostimulant; cardiac; thrombolytic; coagulant; vasotropic;
XX antidiabetic; hypotensive; dermatological; immunosuppressive;
XX antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
XX antithyroid; and antianaemic. The sequences can be used for determining
XX the presence of or predisposition to, or preventing or treating
XX pathological conditions associated with an ORFX-associated disorder. The
XX nucleic acids can be used to express ORFX proteins in gene therapy
XX vectors. The proteins and nucleic acids may be used to treat cancers,
XX proliferative disorders, neurodegenerative disorders, osteoarthritis,
XX graft vs host disease, cardiovascular disease, diabetes mellitus, lupus
XX hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
XX erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
XX bacterial or fungal infection, malaria, autoimmune disorders, asthma,
XX allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
XX nocturnal haemoglobinuria, antiinflammatory disease; to enhance
XX coagulation; to inhibit thrombosis; and as a contraceptive.

XX Sequence 1716 BP; 411 A; 457 C; 446 G; 402 T; 0 other;

Query Match 47.1%; Score 376; DB 21; Length 1716;

Best Local Similarity 68.7%; Pred. No. 6.7e-110;

Matches 533; Conservative 0; Mismatches 240; Indels 3; Gaps 1;

Qy 1 atgcagagctctacagcctgacccagctgccgctatcgcaatgtgttgagccctctgg 60

Dy ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 256 atgcagagcttgctccagatcccaagcccgatcacagatctacggagccctcaag 315

Qy 61 aggattataaagcagagggcctatggaggccatgaggggctgaagctcacagaca 120

Dy ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 316 aaaatctgcgacgaaggctcttgaggcccttgagggcgtcaacgctcatgatcg 375

Qy 121 ggcgcagggcctgcccagccctttatttgcctgctacgaaaagttaaaagacattg 180

Dy ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 376 ggtgcagggccagcccatgcatgtatttgcctgctatgaaacatgaaaggacttta 435

```
Qy 181 agtgatgtaatccaccctgggggcaatagccatattcccaatggtgcggcggtgtgtg 240
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
436 aatgacgttttccaccaccaagaacacagccacctagccaacggatagctgggagtgtg 495
Qy 241 gcaacattactctatgatcagccatgaacacctcggaaagtgttcaagcagagagatcag 300
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
496 gccacctgtccacagatcggttaataatccacagaagtgttgaagcagcgcttcag 555
Qy 301 atgtacaactcaccatcaccacgggtgacagactgtgtacggcgagtggtgcaaaatgaa 360
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
556 atgtacaactcgcagccagcgggtgcgaatcagctgcgtccggcgggtgtggaggaccgag 615
Qy 361 gggggccggggctttaccgcagctacacaccacccagctgaccatgaacgttcccttccaa 420
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
616 ggggtggggccttctaccggagctacacacgcagctgacctgaacatccctctccag 675
Qy 421 gcaattcaactcatgacctatgaattctcgcagagagacacaccccccagagacgtac 480
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
676 tccatccacttcacactatgattctctgcaggagcaggttcaacccccaccggacacctac 735
Qy 481 aaccacaagctccctcctctgtgagcttgcgcaggagctgtagctgcgcagccaca 540
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
736 aaccgcagctcccaacatcctcagcgggctgcgggcccctcgcgcggccgcacag 795
Qy 541 acccactggagctttgcaaacactgctcaacacccagcagagctcttggc---tttgaac 597
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
796 accccctggagctgtgtaagaccttctgaacactcaggagaacgtggccctcgtcgt 855
Qy 598 tcacacattacagagacatatacagggcatggctagtgcttccaggaacgggtatatcaagta 657
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
856 gccaacatcagcggcggtgtcgggtatggccaatgctctccggagcgggtgtaccagctc 915
Qy 658 ggtgggtgaaccgctatttccgaggggtgcagccagagtaatttaccagatccctcc 717
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
916 aacggctgcccgtctactcaaggccatccagcgcgtgtcactcaccagatgccctcc 975
Qy 718 acagccatcgcagtgctgtgtatgattcttcaaatcacttaactcaactaaaggca 773
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
976 accgcaatttctgtgtctatgatttctcaagttaacttctcaccacagcgcca 1031
```

RESULT 15

AAF59920

ID AAF59920 standard; cDNA; 1305 BP.

XX AC AAF59920;

XX DT 04-MAY-2001 (first entry)

XX DE Human mitochondrial solute carrier protein hMSC-o cDNA.

XX DE Human mitochondrial solute carrier protein; hMSC-o; hypothalamus;

XX KW preparation; detection; ss.

XX KW Homo sapiens.

XX OS CN1269409-A.

XX PN 11-OCT-2000.

XX PD 17-MAR-2000; 2000CN-0114958.

XX PF 17-MAR-2000; 2000CN-0114958.

XX PR (SREN-) SOUTHERN RES CENT NAT HUMAN GENE GROUP.

XX PA Zhang X, Gao X, Xiao H;

XX PI WPI: 2001-050544/07.

XX DR P-PSDB; AAB60658.

XX PT New human mitochondrion solute carrier protein and its nucleic acid -

XX XX

PS Claim 1; Page 19-20; 2lpp; Chinese.

XX The invention relates to a novel human mitochondrial solute carrier

CC protein, hMSC-o (AAB60658), and cDNA encoding it (AAF59920). hMSC-o is

CC expressed in normal human hypothalamus tissue. The invention also relates

CC to the preparation of hMSC-o proteins and nucleic acids, and the

CC detection of hMSC-o proteins and nucleic acids in a sample. The present

CC sequence represents cDNA encoding hMSC-o.

XX

SQ Sequence 1305 BP; 292 A; 388 C; 363 G; 262 T; 0 other;

Query Match 45.8%;

Best Local Similarity 68.7%; Pred. No. 1.3e-106;

Matches 533; Conservative 0; Mismatches 239; Indels 4; Gaps 2;

```
Qy 1 atgcagagttctacagcctgaccacgctgcccgctatcgcaatgtgttggaggccctctgg 60
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
172 atgcagagtttgagtcagatcccaagccagtcacaaagtatctcagagccctcaag 231
```

```
Qy 61 aggattataagaacggagggcctatggagccctatggagggcgtgaacgtcacagcaaca 120
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
232 aaaaatcgtcggaccgaaggctctcggagcccttgcgagcgtcaacgtcatgatcatg 291
```

```
Qy 121 ggcgcagggcctgccacgcctttatttgcctgctcacgaaaagttaaaaaagacattg 180
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
292 ggtgcagggcgccgcctatgctgtatttgcctgctatgaaaaacatgaaaaggactta 351
```

```
Qy 181 agtgatgtaataccctggggcgaatagccatatttccaatggtgcggccgggtgtgtg 240
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
352 aatgacgttttccaccaccaagaaacacagccacctagccaacgggatagctggagatg 411
```

```
Qy 241 gcaacattactctatgatgcagccatgaaacccctgcggaagtgttcaagcagagatcag 300
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
412 gccaccctgtccacgctgctgttaataatccacgaagtggtgagcagcgccttcgac 471
```

```
Qy 301 atgtacaactcaccataccaccgggtgcacagactgtgtacggcgagtggtgcaaaatgaa 360
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
472 atgtacaactcgcagcaccgggtcagcaatcagctgcacccggacgggtgtggagcagcg 531
```

```
Qy 361 ggggcccggcctttaccgcagctacacacccagctgacctgaacgttcccttccaa 420
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
532 ggggtggggccttctaccggagctacacacacgagctgacctgaacatcccttccag 591
```

```
Qy 421 gccattcactctatgactatgaattcctgcagagcactttaacccccagagcgttac 480
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
592 tccatccacttccatccactatgattcctgcagagcaggttcaacccccaccggacctac 651
```

```
Qy 481 aacccaagctcccaagctcctctctggagcttgcgcagagcgtgtagctgcgcagccaca 540
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
652 aacccgcagttcccaacatctctcagcggcgtgcggggggcctcgcgcggcgccacg 711
```

```
Qy 541 accccaactggacgtttgcaaaaacactcctcaacacccagagctcttggc---tttgaac 597
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
712 acccccctgacgtctgtaagacccttctgaacactcagagagaacgtggccctctcgtg 771
```

```
Qy 598 tcacacattacagacatatacagcggcatggctagctcctcagagcgggtatatcaagta 657
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
772 gccaaatcagcggcggcgtgtcgggtatggccaatgccttccgcagcgggtgtaccagctc 831
```

```
Qy 658 ggtgggtgaccgcctatttccgagggggtgcagccagcagagtaatttaccagatccctcc 717
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
832 aacggcctg-ccggtacttcaaaagcgtccagcggcgtgtcactcaccagatgccctcc 890
```

```
Qy 718 acagccaatcgcagtggtgtgtgtatgatttcttcaaaatcacctaataactcaataaaggca 773
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
891 accgccatttctgtgtctatgatttctcaagtcttctcaccacagcgcca 946
```

Search completed: September 28, 2002, 08:06:20

Job time: 11575 sec

Query Match 10.2%: Score 81.2; DB 4; Length 1311;
Best Local Similarity 50.2%; Pred. No. 1.1e-16;
Matches 274; Conservative 0; Mismatches 248; Indels 24; Gaps 2;

Qy 50 agggcccttgaggattataaagacggaggcccttatgagggcccatgaggggctgaacg 109
Db 227 AAGCCCTTGGGTGCGATCTAAAGGTGAAGGTCCCGCGGACTTTACCGTGGCATTTGGTG 286

Qy 110 tcacagcaaacaggcaggccctgcccagcccttatttgcctgtacgacgaaagttaa 169
Db 287 CAATGGGTCTCGGTGACGAGCAGCTACGAGGTATTTCTCGTTTACGAGATGTGTA 346

Qy 170 aaagacattgagtgatgaataccacctgggggcaatagccattgccaatggtgcg 229
Db 347 AGGAGACTTTTTCATG-----GTGATCCGAGCAATTCGGTGGCGACGCCGTTT 397

Qy 230 ccgggtgtgtggcaacattactcatgatgcagccatgaaccctgcggaagtgtcgaagc 289
Db 398 CGGGGTGTTCGCGACGGTGGCAAGCGACGCGGTGATTTACGCCGATGGATGTGTGAAC 457

Qy 290 agagatgcagatgtacaaactcaacataccaccgggtgacagactgtgtacgggcagtgt 349
Db 458 AGAGCTTCAGTTGCAGAGCAGTCCGTACAAGGGTGTGTGATTGCGTGAGGAGGTGT 517

Qy 350 gcaaaatgaaggcggcgccctttaccgcagctacacaccagctgaccatgaacg 409
Db 518 TGGTAGAAGAGGATGGCCGCAATTTACGCATCTTATFCGAACAACTGTGTGTCATGAATG 577

Qy 410 ttctttcaagccattcaactcatgatgaacttctgcagggagcaactttaaaccctc 469
Db 578 CCCGTTTACGGCTTCACATTCGCCACATATGAAGCACGAAGAGGGTGTGTGGAGG 637

Qy 470 agagcggtaacaaacccaagctccacgctcct-----ctggagcttgcg 514
Db 638 TGTCCGCGGAGACTGCGAAGCATGAGAAATTTGTAGTGCATGCTGCTGCTGCTG 697

Qy 515 caggagctgtgtgcgcagcccaaccccaactggagcttgcgaacacactgctcaaca 574
Db 698 CTGGAGCTTTGGTGCAGTAGTACCACTCCACTAGATGTTGTCAAAACTCAGTTGCAGT 757

Qy 575 cccagg 580
Db 758 GCCAAG 763

RESULT 2

US-09-068-140A-14
; Sequence 14, Application US/09068140A
; Patent No. 6281409
; GENERAL INFORMATION:
; APPLICANT: Mary Rose Woodhead, Mark Andrew Taylor
; APPLICANT: and Rex Michael Brennan
; TITLE OF INVENTION: Blackcurrant Promoters and Genes
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/068,140A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP96/04807

; FILING DATE: No. 6281409ember 4, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Dinner, Dara L.
; REGISTRATION NUMBER: 33,680
; REFERENCE/DOCKET NUMBER: C70237
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5017
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5150 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Ribes nigrum
; STRAIN: Ben Alder
; US-09-068-140A-14

Query Match 10.2%: Score 81.2; DB 4; Length 5150;
Best Local Similarity 50.2%; Pred. No. 2.2e-16;
Matches 274; Conservative 0; Mismatches 248; Indels 24; Gaps 2;

Qy 50 agggcccttgaggattataaagacggaggcccttatgagggcccatgaggggctgaacg 109
Db 3391 AAGCCCTTGGGTGCGATCTAAAGGTGAAGGTCCCGCGGACTTTACCGTGGCATTTGGTG 3450

Qy 110 tcacagcaaacaggcaggccctgcccagcccttatttgcctgtacgacgaaagttaa 169
Db 3451 CAATGGGTCTCGGTGACGAGCAGCTACGAGGTATTTCTCGTTTACGAGATGTGTA 3510

Qy 170 aaagacattgagtgatgaataccacctgggggcaatagccattgccaatggtgcg 229
Db 3511 AGGAGACTTTTTCATG-----GTGATCCGAGCAATTCGGTGGCGACGCCGTTT 3561

Qy 230 ccgggtgtgtggcaacattactcatgatgcagccatgaaccctgcggaagtgtcgaagc 289
Db 3562 CGGGGTGTTCGCGACGGTGGCAAGCGACGCGGTGATTTACGCCGATGGATGTGTGAAC 3621

Qy 290 agagatgcagatgtacaaactcaacataccaccgggtgacagactgtgtacgggcagtgt 349
Db 3622 AGAGTTGCAGTTGCAGAGCAGTCCGTACAAGGGTGTGTGATTGCGTGAGGAGGTGT 3681

Qy 350 gcaaaatgaaggcggcgccctttaccgcagctacacaccagctgaccatgaacg 409
Db 3682 TGGTAGAAGAGGATTTGGCGCAATTTACGCATCTTATCGAACAACATGTGTGTCATGAATG 3741

Qy 410 ttctttcaagccattcaactcatgatgaacttctgcagggagcaactttaaaccctc 469
Db 3742 CCCGTTTACGGCTTCACATTCGCCACATATGAAGCACGAAGAGGGTGTGTGGAGG 3801

Qy 470 agagcggtaacaaacccaagctccacgctcct-----ctggagcttgcg 514
Db 3802 TGTCCCGGAGACTGCGAAGCATGAGAAATTTGTAGTGCATGCTGCTGCTGCTG 3861

Qy 515 caggagctgtagctgcgcagcccaaccccaactggagcttgcgaacacactgctcaaca 574
Db 3862 CTGGAGCTTTGGTGCAGTAGTACCACTCCACTAGATGTTGTCAAAACTCAGTTGCAGT 3921

Qy 575 cccagg 580
Db 3922 GCCAAG 3927

RESULT 3

US-08-933-750C-61
; Sequence 61, Application US/08933750C
; Patent No. 5932442

RESULT 4
US-09-234-613-61
; Sequence 61, Application US/09234613
; Patent No. 6132973
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti

RESULT 5
US-08-933-750C-68
; Sequence 68, Application US/08933750C
; Patent No. 593242
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga

APPLICANT: Shah, Purvi
APPLICANT: Au-Young, Janice
APPLICANT: Yue, Henry
APPLICANT: Guegler, Karl J.
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
NUMBER OF SEQUENCES: 98
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/933,750C
FILING DATE: September 23, 1997
CLASSIFICATION: 536
PRIOR APPLICATION NUMBER:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0356 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 1643 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear,
IMMEDIATE SOURCE:
LIBRARY: SYNOAT01
CLONE: 724157
US-08-933-750C-68

Query Match 4.58; Score 36; DB 2; Length 1643;
Best Local Similarity 44.0%; Pred. No. 0.073;
Matches 205; Conservative 0; Mismatches 255; Indels 6; Gaps 1;
QY 26 ctgcccctatcgcaatgtgttgagccctctgagattataagaacgagggcctat 85
DB 482 CTACCCCTTCACTGGGACCATGGATGCCCTTCGTGAAGATCGTGAGGCACGAGGGCACCA 541
QY 86 ggagggccatgaggggtgcaagtcacagcaacagcgagggcctgcccacgccccttt 145
DB 542 GGACCCCTCTGGAGCGGCTCCGCCACCCCTGGTGATGACTGTGCCAGCTACCGCCATCT 601
QY 146 atttgctgctacaagaattaaaaagacattgagtgtatccaccctgggggca 205
DB 602 ACTTCACGTGCTGTGACCAACTGAAAGCCCTTCCTGTGTGTGTCGAGGCC-----TGACCT 655
QY 206 atagccatattgccattggtgcgcccgggtgtgtgcaacattacttcattgatgcacca 265
DB 656 CTGACCTCTACGCACCCATGGTGTGCTGGCGGCTGGCGGCTGGGACCGTGACTGTGA 715
QY 266 tgaacctcggaagtgttcaagcagaggtatgcagattgtacaactcaccataccacggg 325
DB 716 TCAGCCCCCTGGAGCTTATATGCGGACAAAGCTGACGGCTCAGAGCTGTGTGTCACCGGAGC 775
QY 326 tgacagactgtgtacggcgagtggtgcaaatgaagggcgccggtttttaccgagct 385
DB 776 TGGGTGCTCTGTGTCGAACCTGCAGTGCTCAGGGTGGCTGGCGTCACTGTGTGGTGGCT 835

QY 386 acaccaccagctgaccatgaacgttcttttcccaagcattcacttcacttcacctatgaat 445
DB 836 GGGCCCCCACTGCCCTTCGAGATGTGCCCTTCTCAGCCCTGTACTGGTTCAACTATGAGC 895
QY 446 tctgcagagagcactttaacccccagagagcggtacaaaccaagctc 491
DB 896 TGGTGAAGAGCTGGCTCAATGGGCTCAGGCCGAAGGACCAGACATTC 941
RESULT 6
US-09-234-613-68
Sequence 68, Application US/09234613
Patent No. 6132973
GENERAL INFORMATION:
APPLICANT: Lal, Preeti
APPLICANT: Hillman, Jennifer L.
APPLICANT: Bandman, Olga
APPLICANT: Shah, Purvi
APPLICANT: Au-Young, Janice
APPLICANT: Yue, Henry
APPLICANT: Guegler, Karl J.
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
NUMBER OF SEQUENCES: 98
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/234,613
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/933,750
FILING DATE: September 23, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0356 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 1643 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: SYNOAT01
CLONE: 724157
US-09-234-613-68

Query Match 4.58; Score 36; DB 3; Length 1643;
Best Local Similarity 44.0%; Pred. No. 0.073;
Matches 205; Conservative 0; Mismatches 255; Indels 6; Gaps 1;
QY 26 ctgcccctatcgcaatgtgttgagccctctgagattataagaacgagggcctat 85
DB 482 CTACCCCTTCACTGGGACCATGGATGCCCTTCGTGAAGATCGTGAGGCACGAGGGCACCA 541
QY 86 ggagggccatgaggggtgcaagtcacagcaacagcgagggcctgcccacgccccttt 145

Db 542 GGACCTCTGGAGCGCCCTCCCGCCACCCCTGGTGATGACTGTGCCAGCTACCGCCATCT 601
Qy 146 attttgcctgtacgaaagtttaaaagacatttgatgtatgaatccaccctggggca 205
Db 602 ACTTCACCTGCTATGACCAACTGAGGCCCTTCTGTGTGGTCGAGCCC-----TGACCT 655
Qy 206 atagccatattgccaattggtgcgcgcgggtgtgtgcaacatttaattcatgatgcagcca 265
Db 656 CTGACCTCTACGACCCATGCTGTGGTGGCGCTGGCGCTGGCGACCGGTGACTGTGA 715
Qy 266 tgaacccctgcgaagtgtcaagcagagagatgcagatgtataactcacatccaccggg 325
Db 716 TCAGCCCCCTGAGAGCTATGCGGCAAAAGCTGACGGCTGAGTGTGTCACCGGAGC 775
Qy 326 tgacagactgttcagggcagtggtgcaaaatgaaggggcgccgttttaccgcagct 385
Db 776 TGGGTGCTGTGTTGCACTGAGTGGCTCAGGTGGCTGCGGCTCACTGTGGCTGGGT 835
Qy 386 acaccaccagctgaccatgaacgtttcttccaaagccattcaattcatgacctatgaat 445
Db 836 GGGGCCCCACTGCCCTTCGAGATGTGCCCTTCTCAGCCCTGTACTGGTTCAACTATGAGC 895
Qy 446 tcttcagagagcactttaacccccagagagcaggtacaaacccaagctc 491
Db 896 TGGTGAAGAGCTGGCTCAATGGGCTCAGGCGGAGGACGAGACTTC 941

RESULT 7

US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F1s
US-08-232-463-14

Query Match 4.4%; Score 35; DB 1; Length 7218;
Best Local Similarity 4.1%; Pred. No. 0.33;
Matches 14; Conservative 183; Mismatches 148; Indels 0; Gaps 0;

Qy 59 ggaagattataagaacgagggccttatgagcccatgaggggctgaacgtcacagcaa 118
Db 1383 RRR 1324
Qy 119 caggcgagggcctcccccagcccttattttgctgtctacgaaagttaaaaaagacat 178
Db 1323 RRR 1264
Qy 179 tgagtgtatgaataccaccctgggggcaatagcattatgccaatggtgcgcgggtgtg 238
Db 1263 RRR 1204
Qy 239 tggcaacattactcatgagccatgaaccctgcggaagtgtcaagcagagagatgc 298
Db 1203 RRR 1144
Qy 299 agatgtacaaactcacataccaccgggtgacagactgtgtacggcgagtggtgcaaaatg 358
Db 1143 RRR 1084
Qy 359 aagggggcgggcctttaccgcagctacacccaccagctgacca 403
Db 1083 RRR 1039

RESULT 8

US-08-937-466-5
; Sequence 5, Application US/08937466
; Patent No. 5846779
; GENERAL INFORMATION:
; APPLICANT: Zhang, Ning
; APPLICANT: Amaral, M. Catherine
; APPLICANT: Chen, Jin-Long
; TITLE OF INVENTION: UCP3 Genes
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/937,466
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: T97-009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1777 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear

; MOLECULE TYPE: CDNA
US-08-937-466-5

Query Match	4.1%	Score 32.4;	DB 2;	Length 1777;
Best Local Similarity	44.1%;	Pred. No. 1.1;		
Matches 135;	Conservative 0;	Mismatches 171;	Indels 0;	Gaps 0;

Qy	244	acattactcatgatgcagccatgaacctcggaagtggccaagcagagatgcagatg	303
Db	330	ACTTTTCCCTGGACACGCCAAGTCCGTCTGCAGATCCCAAGGGGAGAACCCAGGGGCT	389
Qy	304	tacaactcacataccaccgggtgacagactgtgtacggggcagttggcgaatatgaaggg	363
Db	390	CAGAGCGTGCAGTACC CGGGTGTGCTGGTACCATCCTGACTATGTGTCGCACAGAGGCT	449
Qy	364	gcgggggcccctttaccgcaggtacacacaccagctgaacatgaacgttcccttccaagcc	423
Db	450	CCCCGCGACGCCCTACAGCGGGACTGTGCTCGCTGGCTGCACCGCCACAGATGAGTGTTCGCTTCC	509
Qy	424	attcacitcatgacctatgaattcctgcagagacatttaaccgccagagcgttacaac	483
Db	510	ATT CGAATTGGCCCTCTAGACTCTGTCAAGCAGTTCTACACCCCAAGGGACGGACAC	569
Qy	484	ccaagctcccacgtccctctctctggagcttgcgcaggagctgtgattgcgcgcagccacaacc	543
Db	570	TCCAGCTCGCCATCAGGATTCTGGCAGGCTGCAGCAGGAGGCCATGTCAGTGCCTG	629
Qy	544	ccactg	549
Db	630	GCCCAG	635

```

RESULT          9
US-09-172-528-5
; Sequence 5, Application US/09172528
; Patent No. 5952469
; GENERAL INFORMATION:
; APPLICANT: Zhang, Ning
; APPLICANT: Amaral, M. Catherine
; APPLICANT: Chen, Jin-Long
; TITLE OF INVENTION: UCP3 Genes
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/172,528
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/937,466
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: T97-009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1777 base pairs
; TYPE: nucleic acid

```

```

; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-09-172-528-5

Query Match          4.1%; Score 32.4; DB 2; Length 1777;
Best Local Similarity 44.1%; Pred. No. 1,1;
Matches 135; Conservative 0; Mismatches 171; Indels 0; Gaps

Qy 244 acattactcatgatgcagccatgaaccctgcggaagtgtgccaagcagagagatgcagatg 303
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 330 ACTTTTCCCTGGACACCGCCAAAGTCCGCTCGATCCAAAGGGAGAACCCAGGGGCT 389
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 304 tacaactcacatataccacgggtgcagactgtgtacgggcagtgtgccaatgaagg 363
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 390 CAGAGCGTGCAGTACCCGGGTGCTGGGTACCATCTGACTATGTTGGCCACAGAGGGT 449
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 364 gccggggcctttaccgcagctacaccccagctgaccatgaacgttcctttcccaagcc 423
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 450 CCCCgcAGCCCCFACAGCGGACTGGTGGCTGGCTGCACGCCAGATGAGTTTGGCTCC 509
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 424 attcaactatgacctatgaattcctctgcagagacactttaacccccagagacggtacaac 483
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 510 ATTCCGAATTGGCTCTACGACTCTGTCAACGAGTTCTACACCCCAAGGGAGCGGACCAC 569
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 484 ccaagctcccaacgttcctctctggagcttgcgagagactgtagctgcccgagccacaacc 543
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 570 TCCAGCGTCGCCATCAGGATTCTGGCAGGCTGCACGACAGGACCATGGCAGTGACCTGC 629
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 544 ccacty 549
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 630 GCCCAG 635
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 10
US-09-318-199-5
: Sequence 5, Application US/09318199
: Patent No. 6025469
: GENERAL INFORMATION:
: APPLICANT: Zhang, Ning
: APPLICANT: Amaral, M. Catherine
: APPLICANT: Chen, Jin-Long
: TITLE OF INVENTION: OCP3 Genes
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
: STREET: 75 DENISE DRIVE
: CITY: HILLSBOROUGH
: STATE: CALIFORNIA
: COUNTRY: USA
: ZIP: 94010
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/318,199
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/937,466
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: OSMAN, RICHARD A
: REGISTRATION NUMBER: 36,627
: REFERENCE/DOCKET NUMBER: T97-009
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (650) 343-4341
: TELEFAX: (650) 343-4342
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:

```

```
; ; LENGTH: 1777 base pairs
; ; TYPE: nucleic acid
; ; STRANDEDNESS: double
; ; TOPOLOGY: linear
; ; MOLECULE TYPE: cDNA
US-09-318-199-5

Query Match          4.1%; Score 32.4; DB 3; Length 1777;
Best Local Similarity 44.1%; Pred. No. 1.1;
Matches 135; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

Qy 244 acattacttcattgatgcagccatgaacccctgcgggaagtgtcaaacgagaggtgcagatg 303
Db 330 ACTTTTCCCTTGGACACCGCCCAAGGTCCGTCGACAGATCCCAAGGGGAGAACCCAGGGGCT 389
Qy 304 tacaactcaccatcacccgggtgcagactgtgtacgggcagtggtgcaaaatgaaagg 363
Db 390 CAGACGGTGCAGTACCGGGTGTGTGGGTACCATCCCTGACTATGTTGGTGGCACAGAGGT 449
Qy 364 gccggggccttttaaccgcagctacacccaccagctgaccatgaacgttcccttccaaagg 423
Db 450 CCCCAGCCCTACACGGGACTGTGCTGCTGGCTGCACCCGACATGATGTTTGCCTCC 509
Qy 424 attacttcattgaactatgaattcctgcaggagcactttaacccccagagcggtacaac 483
Db 510 ATTCAATTGGCCTTACGACTCTGTCAAGCAGTTCTACACCCCAAGGGAGCGGACCAC 569
Qy 484 ccaagctcccagctctctggaacttgccagagctgtagctgcccagccacaacc 543
Db 570 TCCACGTCGCCATCAGGATTCTGCGAGGCTGCACGACAGGAGCCATGGCAGTGACCTGC 629
Qy 544 ccaactg 549
Db 630 GCCCAG 635

RESULT 11
US-09-503-579-5
; Sequence 5, Application US/09503579
; Patent No. 6248561
; GENERAL INFORMATION:
; APPLICANT: Zhang, Ning
; APPLICANT: Amaral, M. Catherine
; APPLICANT: Chen, Jin-Long
; TITLE OF INVENTION: UCP3 Genes
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE:
; PRIOR APPLICATION NUMBER: US/09/503,579
; CLASSIFICATION:
; APPLICATION NUMBER: US/08/937,466
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: T97-009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342

; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1777 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-503-579-5

Query Match          4.1%; Score 32.4; DB 4; Length 1777;
Best Local Similarity 44.1%; Pred. No. 1.1;
Matches 135; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

Qy 244 acattacttcattgatgcagccatgaacccctgcgggaagtgtcaaacgagaggtgcagatg 303
Db 330 ACTTTTCCCTTGGACACCGCCCAAGGTCCGTCGACAGATCCCAAGGGGAGAACCCAGGGGCT 389
Qy 304 tacaactcaccatcacccgggtgcagactgtgtacgggcagtggtgcaaaatgaaagg 363
Db 390 CAGACGGTGCAGTACCGGGTGTGTGGGTACCATCCCTGACTATGTTGGTGGCACAGAGGT 449
Qy 364 gccggggccttttaaccgcagctacacccaccagctgaccatgaacgttcccttccaaagg 423
Db 450 CCCCAGCCCTACACGGGACTGTGCTGCTGGCTGCACCCGACATGATGTTTGCCTCC 509
Qy 424 attacttcattgaactatgaattcctgcaggagcactttaacccccagagcggtacaac 483
Db 510 ATTCAATTGGCCTTACGACTCTGTCAAGCAGTTCTACACCCCAAGGGAGCGGACCAC 569
Qy 484 ccaagctcccagctctctggaacttgccagagctgtagctgcccagccacaacc 543
Db 570 TCCACGTCGCCATCAGGATTCTGCGAGGCTGCACGACAGGAGCCATGGCAGTGACCTGC 629
Qy 544 ccaactg 549
Db 630 GCCCAG 635

RESULT 12
US-08-937-466-3
; Sequence 3, Application US/08937466
; Patent No. 5846779
; GENERAL INFORMATION:
; APPLICANT: Zhang, Ning
; APPLICANT: Amaral, M. Catherine
; APPLICANT: Chen, Jin-Long
; TITLE OF INVENTION: UCP3 Genes
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE:
; APPLICATION NUMBER: US/08/937,466
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: T97-009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 3:
```


TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1949 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cdna
US-09-318-199-3

Query Match 4.1%; Score 32.4; DB 3; Length 1949;
Best Local Similarity 44.1%; Pred. No. 1.2;
Matches 135; Conservative 0; Mismatches 171; Indels 0; Gaps 0;
QY 244 acattacttcattgatgcagccatgaacccctgcgggaagtgtcaagcagaggatgcagatg 303
DB 330 ACTTTTCCCTGGACACCCGCCAAGGTCCGTCTGCAGATCCAAAGGGGAGAAACCCAGGGGCT 389
QY 304 tacaactcaccatacaccgcgggtgacagactgtgtacgggcagtggtgcaaaatgaagg 363
DB 390 CAGACGTCGACTACCGGGTGTGTGGTACCATCTCTGACTATGTCGGCACACAGGGT 449
QY 364 gccggggcctttaccgcagctacacccaccagctgaccatgaacgttcttcccaagcc 423
DB 450 CCCCCAGCCCCCTACAGCGGACTGGTCGCTGGCCCTGCACCGCCAGATGAGTTTGCCTCC 509
QY 424 attacttcattgatgcagccatgaacccctgcgggaagtgtcaagcagaggatgcagatg 483
DB 510 APTCGAATGGGCTCTACGACTCTGTCAAGCAGTTCTACACCCCAAGGGAGCGGACCCAC 569
QY 484 ccaagctcccagctctctctgagcttgcgcaggagctgtagctgcgcagcccaaac 543
DB 570 TCCAGCGTCGCCATCAGGATTCGCGAGCTGCACGACAGGAGCCATGGCAGTGACCTGC 629
QY 544 ccaactg 549
DB 630 GCCCAG 635

RESULT 15
US-09-503-579-3
Sequence 3, Application US/09503579
Patent No. 6248561
GENERAL INFORMATION:
APPLICANT: Zhang, Ning
APPLICANT: Amaral, M. Catherine
APPLICANT: Chen, Jin-Long
TITLE OF INVENTION: UCP3 Genes
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 75 DENISE DRIVE
CITY: HILLSBOROUGH
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94010
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/503,579
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/937,466
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A

REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: T97-009
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1949 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cdna
US-09-503-579-3
Query Match 4.1%; Score 32.4; DB 4; Length 1949;
Best Local Similarity 44.1%; Pred. No. 1.2;
Matches 135; Conservative 0; Mismatches 171; Indels 0; Gaps 0;
QY 244 acattacttcattgatgcagccatgaacccctgcgggaagtgtcaagcagaggatgcagatg 303
DB 330 ACTTTTCCCTGGACACCCGCCAAGGTCCGTCTGCAGATCCAAAGGGGAGAAACCCAGGGGCT 389
QY 304 tacaactcaccatacaccgcgggtgacagactgtgtacgggcagtggtgcaaaatgaagg 363
DB 390 CAGACGTCGACTACCGGGTGTGTGGTACCATCTCTGACTATGTCGGCACACAGGGT 449
QY 364 gccggggcctttaccgcagctacacccaccagctgaccatgaacgttcttcccaagcc 423
DB 450 CCCCCAGCCCCCTACAGCGGACTGGTCGCTGGCCCTGCACCGCCAGATGAGTTTGCCTCC 509
QY 424 attacttcattgatgcagccatgaacccctgcgggaagtgtcaagcagaggatgcagatg 483
DB 510 APTCGAATGGGCTCTACGACTCTGTCAAGCAGTTCTACACCCCAAGGGAGCGGACCCAC 569
QY 484 ccaagctcccagctctctctgagcttgcgcaggagctgtagctgcgcagcccaaac 543
DB 570 TCCAGCGTCGCCATCAGGATTCGCGAGCTGCACGACAGGAGCCATGGCAGTGACCTGC 629
QY 544 ccaactg 549
DB 630 GCCCAG 635

Search completed: September 28, 2002, 07:55:54
Job time: 11239 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2002, 06:22:41 ; Search time 4881.77 Seconds
(without alignments)
2206.284 Million cell updates/sec

Title: US-09-870-113-7
Perfect score: 798
Sequence: 1 atgcagagctacagcctga.....agtggaggctggcaagtga 798

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues
Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	628.8	78.8	747	10	BI827322
2	587	73.6	826	10	BI254253
3	586	73.4	986	10	BC331197
4	532.2	66.7	576	10	BI057218
5	517	64.8	927	10	BI0972351
6	507.4	63.6	908	10	BC697119
7	494	61.9	855	10	BI412175
8	473	59.3	808	10	BC918082
9	472.6	59.2	531	9	AW915209
10	470.6	59.0	563	10	BF193023
11	468	58.6	581	9	AW341177
12	453.6	56.8	830	10	BF143133
13	453.4	56.8	578	10	BM483098
14	437.2	54.8	888	9	AL530804
15	435	54.5	451	10	BF515180
16	433.8	54.4	529	9	AA431276
17	433	54.3	527	9	AI797090

18	432	54.1	526	9	AI797100
19	429.6	53.8	741	9	AL563845
20	427.2	53.5	709	9	AJ397503
21	424.2	53.2	560	10	BI794616
22	423.8	53.1	1131	10	BG295496
23	421	52.8	681	9	AV704087
24	420.2	52.7	645	10	BG964218
25	417.8	52.4	483	9	AA061624
26	405.6	50.8	563	10	BF078769
27	404	50.6	598	10	BE913718
28	398.4	49.9	541	9	AA234031
29	397.6	49.8	574	10	BI185580
30	393.6	49.3	453	10	BF916224
31	391	49.0	488	9	AI797101
32	387.8	48.6	1295	11	AK015770
33	387	48.5	480	10	BI045863
34	387	48.5	718	10	BJ041453
35	384.2	48.1	501	9	AW211366
36	383.2	48.0	423	10	BI339456
37	382.8	48.0	425	10	BI304641
38	381.6	47.8	848	9	AI743110
39	371	46.5	434	10	W53634
40	368.2	46.1	491	10	BF076673
41	363.8	45.6	1209	11	AK006155
42	355.4	44.5	732	9	AI208913
43	350	43.9	466	10	BM258580
44	346.4	43.4	586	10	BG982657
45	332.2	41.6	490	10	BG055067

ALIGNMENTS

RESULT 1

BI827322

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BI827322 747 bp mRNA linear EST 04-OCT-2001
603077989F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5169694 5',
mRNA sequence.
BI827322 GI:15938872
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/
1 (bases 1 to 747)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1422 row: a column: 23
High quality sequence stop: 743.
Location/Qualifiers
1. .747
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5169694"
/clone_lib="NIH_MGC_119"
/tissue_type="medulla"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site:1: NotI;
Site:2: EcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb, insert size range

FEATURES
source

0.9-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 013. Note: this is a NIH_MGC Library.

BASE COUNT 188 a 206 c 201 g 151 t 1 others

Query Match 78.8%; Score 628.8; DB 10; Length 747;
Best Local Similarity 99.4%; Pred. No. 3.9e-172;
Matches 652; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

Qy 1 atgcagagctacagctgaccagctgcccgcctatcgcaatgtgtgagccctctgg 60
Db 84 ATGCAGAGCTACAGCTGACCGTGCAGCTGCCGCTATCGCAATGTGTGAGGCCCTCTGG 143
Qy 61 aggattataagAACAGGAGGCGCTATGGAGGCCCATGAGGGGCTGAACGTCACAGCAACA 120
Db 144 AGGATTATAAGAACAGGAGGCGCTATGGAGGCCCATGAGGGGCTGAACGTCACAGCAACA 203
Qy 121 ggcagggcgctgccagccctttatttgcctgctctacgaagaatttaaaagacattg 180
Db 204 GCGCAGGGCGCTGCCAGCCCTTTATTTCCTGCTAGCAAAAGTTAAAAAGACATTTG 263
Qy 181 agtgatglaatccaccctgggggcaatagccatattgccaatgtgcgccgggtgtg 240
Db 264 AGTGATGTAATCCACCTGGGGGCAATAGCCATATTGCCAATGTGTGGCGCGGTGTGG 323
Qy 241 gcaattacttcatgctgagccatgaacctcggaagtgtgtcaagcagaggtcag 300
Db 324 GCAACATTACTTCATGATGTCAGCCATGAACCTTGGGAAGTGGTCAAGCAGAGGATGCAG 383
Qy 301 atgtacaactcaaccaccgggtgacagactgtgacggcagtggtgcaaatgaa 360
Db 384 ATGTACACTCAACATACCACCGGTGTACAGACTGTGTACGGGCAAGTGTGCAAAATGAA 443
Qy 361 ggggcccgggcttttacgcagctacaccacccagctgacctgaacgttctcttccaa 420
Db 444 GGGCGCGGGCGCTTTTACCGCAGCTACACCACCCAGCTGACCATGAACGTTCTTTCCAA 503
Qy 421 gccattacttcatgactatgaattctctgagggagcactttaacccacagagcgttac 480
Db 504 GGCATTCACTTCACTGACCTAATTCCTGTCAGAGAGCACTTTTAACCCCCAGAGCGGTAC 563
Qy 481 ascccaagctccacgtctctctgtgagcttgcgagagctgtgagctgctgcttgcaccca 540
Db 564 AACCCAAAGTCCCGCTCTCTGTGAGCTTGGCAGAGAGCTGTAGCTGCCAGCCACCA 623
Qy 541 acccactggagct-ttgcaaacactgctcaacacccagggagctctgttggactc 599
Db 624 GACCCACTGGAGCTTNTTGCAAAACACTGCTCAACACCCAGGAGTCTTGGCTTTGAAC 683
Qy 600 acacattacagga-catatcacagcagctggttagtgccttcagcagcgttatataca 654
Db 684 ACACATTACAGGACCATATACAGGCATGGCTAGTGCCTTTCAGGAGCGGTATATACAA 739

RESULT 2
B1254253 826 bp mRNA linear EST 17-JUL-2001
LOCUS 602974729F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5114190 5',
DEFINITION mRNA sequence.

ACCESSION B1254253
VERSION B1254253.1 GI:14806485
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 826)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L14M11277 row: i column: 07
High quality sequence stop: 825.

FEATURES

source
1. 826
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5114190"
/clone_lib="NIH_MGC_12"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.4 kb. Library prepared by Life Technologies."

BASE COUNT 211 a 233 c 208 g 174 t
ORIGIN

Query Match 73.6%; Score 587; DB 10; Length 826;
Best Local Similarity 99.8%; Pred. No. 6.1e-160;
Matches 598; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 200 ggggcaatagccatattgccaatgtgcccgggtgtggaacattactcatgatg 259
Db 1 GGGGCAATAGCCATA-TGCCAATGTGCGCGGGGTGTGGCAACATTACTTTCATGATG 59
Qy 260 cagccatgaacctcggaagtgtgcaagcagagatgcagatgtacaactcaccatacc 319
Db 60 CAGCCATGAACCTCGGAAGTGGTCAAGCAGAGATGCAGATGTACAACCTCACCATACC 119
Qy 320 accgggtgacagactgtgtacgggcagtggtgcaaaaagaagggcgcccttttaac 379
Db 120 ACCGGGTGACAGACTGTGTACGGGCAGTGTGGCAAAATGAAGGGGCGGGGCTTTTAC 179
Qy 380 gcagctacaccaccagctgacctgaacgttcttccaaagccattcaactcatgacct 439
Db 180 GCAGCTACACCACCAGCTGACCATGAAGTTCCTTTCACAGCCATTCACTTCATGACCT 239
Qy 440 atgaattcctgcaggagcactttaacccccagagacggttacaacccagctccccagtc 499
Db 240 ATGAATTCTTCAGGAGCACTTTAAACCCCAAGAGCGGTACACCCAGCTCCCAAGTCC 299
Qy 500 tctctggagcttgcaggagctgtagctgcgagcagcacaacccactggagctttgca 559
Db 300 TCTCTGGAGCTTGGCGCAGGAGCTGTAGCTGCCGAGCCACACACCCCTGAGCTTTC 359
Qy 560 aaacactcacaacacccagagctcttggcttgaactcacacattacagacatatca 619
Db 360 AAACACTGCTCAACACCACAGGAGTCTTGGCTTTGAACATACACATACAGACATATCA 419
Qy 620 caggcatggttagtgccttccagagcgttatatacagtaggtggtgagccgctatttcc 679
Db 420 CAGGATGGCTAGTGCCITTCAGGACGGTATATCAAGTAGGTGGGTGAGCCGCTATTTC 479
Qy 680 gaggggtgcagccagagtaatttaccagatccccctcacagccatcgcatggtctgt 739
Db 480 GAGGGGTGCAGGCCAGAGTAATTTTACCAGATCCCCCTCCACAGCCATCGCATGCTGT 539
Qy 740 atgagttctcaaatcacatcaatacctaaagcaagaagtgaggcgtggaagtga 798
Db 540 ATGAGTTCTTCAATACCTTAATCACTAAAGCAAGAGAGTGGAGGGCTGCCAAGTGA 598

RESULT 3

```

BG331197
LOCUS      BG331197      986 bp      mRNA      linear      EST 27-FEB-2001
DEFINITION 602431907F1 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:4549801 5',
            mRNA sequence.
ACCESSION  BG331197
VERSION     BG331197.1 GI:13137635
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 986)
AUTHORS     NIH-MGC http://mgi.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/BTP/Gazdar
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: L1CM1241 row: m column: 02
            High quality sequence stop: 387.
            Location/Qualifiers
              1..986
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:4549801"
                /tissue_type="large cell carcinoma"
                /lab_host="DH10B (phage-resistant)"
                /note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
                EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                into EcoRI/XhoI sites using the following 5' adaptor:
                GGCACGAG(G). Library constructed by Ling Hong in the
                laboratory of Gerald M. Rubin (University of California,
                Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                Superscript II RT (Life Technologies). Note: this is a
                NIH_MGC Library."
BASE COUNT  268 a 277 c 248 g 193 t
ORIGIN
Query Match 73.4%; Score 586; DB 10; Length 986;
Best Local Similarity 97.5%; Pred. No. 1.3e-159;
Matches 595; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 187 gtaatccaccctgggggcaatagccatattgccaatggtgcggcggtgtgtggcaaca 246
|
|
|
Db 2 GTAATCCACCCTGGGGGCAATAGCCATATTGCCAATGTGCGCGGCTGTGTGCAACA 61
|
|
|
Qy 247 ttacttcgatgatgagccatgaacctcgcgaagtgttcacgagagagatgcagatgtac 306
|
|
|
Db 62 TTACTTTCATGTATGACGCATGAACCCCTCGGAAGTGGTCAAGAGAGAGATGCAGATGTAC 121
|
|
|
Qy 307 aactcaccataccaccgggtgacagactgtgacgggcagtggtgacaaatgaagggcc 366
|
|
|
Db 122 AACTCACCATACCACCGGGTGACAGACTGTGTACGGGCGAGTGTGGCAAAATGAAGGGGCC 181
|
|
|
Qy 367 ggggcccctttaccgcagctacaccaccagctgacatgaaccttccttccaaagccatt 426
|
|
|
Db 182 GGGGCTTTTACCAGAGCTACACACCAGCTGACCATGACCATGACGTTCCCTTCCAGCCATT 241
|
|
|
Qy 427 cacttcgatcctatgaattcctgcagagacaccttaacccccagagacggtacaaacca 486
|
|
|
Db 242 CACTTCATGACCTATGAATTCCTGCAGGAGCAGCTTAAACCCCCAGAGACGGGTACAACCCA 301
|
|
|
Qy 487 agctcccaagctcctctctgagcttgacagagctgtagtgcgcagcagccacaaccca 546
|
|
|
Db 302 AGCTCCCAAGCTCCTCTCTGGAGCTTGGCAGGAGCTGTAGCTGCCGCGAGCCACAACCCA 361
|
|
|

```

```

Qy 547 ctggagcgtttgcaaacactgctcaacccccaggagctcttgcttgaactcacacatt 606
|
|
|
Db 362 CTGGAGCTTTGCAAAACACTGCTCAACCCAGGAGTCTTGGCTTTGAACTCACACATT 421
|
|
|
Qy 607 acaggacatatcacaggcatggctagtgcttcacgagcaggtatatacaagtgtgggtg 666
|
|
|
Db 422 ACAGGACATATCACAGGCATGGCTAGTGCTTTCAGGACGGTATATCAAGTAGTGGGTG 481
|
|
|
Qy 667 accgcctattccgaggggtgcagccagagtaattaccagatccctccacacaccatc 726
|
|
|
Db 482 ACCGCTATTTCCGACGGGTGCGCCAGAGTAATTTACAGATCCCTTCCACAGCCATC 541
|
|
|
Qy 727 gcatgctgtgtatgagttcttcaaatcacctaaatcactaaaggcaagaagtgagg 786
|
|
|
Db 542 GCATGCTGTGTATGAGTTCTTCAATACCTAATCACTAAAGGCAAGACACTGGAGG 601
|
|
|
Qy 787 gctggcaagt 796
|
|
|
Db 602 GCTGGCAAGT 611
|
|
|
RESULT 4
BIOS7218/c
LOCUS      QV3-GN0338-230201-558-a01 GN0338 Homo sapiens cDNA, mRNA sequence.
DEFINITION QV3-GN0338-230201-558-a01 GN0338 Homo sapiens cDNA, mRNA sequence.
ACCESSION  BI057218
VERSION     BI057218.1 GI:14464748
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 576)
AUTHORS     Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
            Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
            Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
            20202663
            Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?tl-QV3&t2-QV3-GN0338-
            230201-558-a01&t3=2001-02-23&t4=1)
            Seg primer: puc 18 forward
            High quality sequence start: 18
            High quality sequence stop: 575.
            Location/Qualifiers
              1..576
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone_lib="GN0338"
                /dev_stage="Adult"
                /note="Organ: placenta_normal; Vector: puc18; Site_1: SmaI
                ; Site_2: SmaI; A mini-library was made by cloning
                products derived from ORESTES PCR (U.S. Letters Patent
                application No. 196.716 - Ludwig Institute for Cancer
                Research) profiles into the puc 18 vector. Reverse
                transcription of tissue mRNA and cDNA amplification were
                performed under low stringency conditions."
BASE COUNT  129 a 141 c 162 g 144 t

```


REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA
Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10699 row: o column: 10
High quality sequence stop: 878.

BASE COUNT 218 a 236 c 234 g 220 t

ORIGIN

Query Match 63.6%; Score 507.4; DB 10; Length 908;
Best Local Similarity 98.8%; Pred. No. 9.8e-137;
Matches 574; Conservative 0; Mismatches 1; Indels 6; Gaps 6;

QY 223 gttgcccgggtgtgtgcaacattactatgatgagccatgaacccctgcgaagtg 282
|||||
Db 244 ggtggccgggtgtgtgcaacattactatgatgagccatgaacccctgcgaagtg 303
|||||
QY 283 gtcaagcagagatgcagatgtacaactcaaccacccgggtgacagctgtgacgg 342
|||||
Db 304 gtcaagcagagatgcagatgtacaactcaaccacccgggtgacagctgtgacgg 363
|||||
QY 343 gcaagtgtgcaaaatgaagggccggggccctttaccgcagctacaccacccagctgacc 402
|||||
Db 364 gcagtggtgcaaaatgaagggccggggccctttaccgcagctacaccacccagctgacc 423
|||||
QY 403 atgaacgttcccttccaagccattcaactcatgacctatgaattcctgcagagacattt 462
|||||
Db 424 atgaacgttcccttccaagccattcaactcatgacctatgaattcctgcagagacattt 483
|||||
QY 463 a-accgccagagacggtacaaccccaagctccacgctcctctctgagcttgcgaggagc 521
|||||
Db 484 ataccgccagagacggtacaaccccaagctccacgctcctctctgagcttgcgaggagc 543
|||||
QY 522 tctagctgcgcagccacaccccccactgacgctttgcaaaacactgctcaaca-cccagg 580
|||||
Db 544 tctagctgcgcagc-CACAACCCCACTGGACGCTTTGCAAAACACTGCTCAACAGCCCAAG 602
|||||
QY 581 agtccttgcttgactcaacacac-ttacagacatatcaaggcattgctagtgccttc 639
|||||
Db 603 agtccttgcttgactcaacacac-ttacagacatatcaaggcattgctagtgccttc 662
|||||
QY 640 a-ggacgggtatcatcaagtaggtgggtgaccgccttattccaggggtgcagggccagagt 698
|||||
Db 663 agggacggtatcatcaagtaggtgggtgaccgccttattccaggggtgcagggccagagt 722
|||||
QY 699 aattaccagatccctccacagccatgcagctggtgtgtgtatgatgttctcaaatatcct 758
|||||
Db 723 aatttaccagatccctccacagccatgcagctggtgtgtgtatgatgttctcaaatatcct 782
|||||
QY 759 aatcactaaaaaggaagaag-agtggagggtgtgcaagtga 798

Db 783 AATCACTAAAGGCAAGAAAGTGGAGGCTGGCAAGTGA 823
|||||

RESULT 7
BI412175
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
JOURNAL
COMMENT

BI412175 602965377F1 NCI_CGAP_Lu33 Mus musculus cDNA clone linear EST 14-AUG-2001
mRNA sequence.
BI412175
BI412175.1 GI:15173098
EST.
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11294 row: j column: 20
High quality sequence start: 7
High quality sequence stop: 809.

BASE COUNT 209 a 233 c 241 g 172 t

ORIGIN

Query Match 61.9%; Score 494; DB 10; Length 855;
Best Local Similarity 87.9%; Pred. No. 7.6e-133;
Matches 675; Conservative 0; Mismatches 80; Indels 13; Gaps 12;

QY 1 atgcagagctcacagcctgacccag-ctgccgcctatgcgaatgtgttgaggccctctg 59
|||||
Db 86 atgcagagcctacagcctgacccagacccgcctatgcgaacgctgttgaggcctctctg 145
|||||
QY 60 gaggattaaagcagggagcctatggagggcccatgagggg-gctgaacgtcacagcaa 118
|||||
Db 146 gagatcatcagacagagaggcctgttgaggcccatgagggtgctgacgctcagacaa 205
|||||
QY 119 caggcgagggcc-tgccacagccct-ttattttgcctgctacgaaaaatctctctctctct 176
|||||
Db 206 caggcgagggccgtgccacagccctctgatttttgcctgctacgaaaaatctctctctct 265
|||||
QY 177 a-ttgagtgtatgaatccaccctgggggccaatagccaattgtgcgcccgggt 235

Kerlavage,A.R. and Adams,M.D.
Rat Genome Project: Generation of a Rat EST (REST) Catalog & Rat
Gene Index
Unpublished (1998)
Contact: Lee, NH
The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529
Fax: (301)-838-0208
Email: nhlee@tigr.org
This clone is available through the ATCC, contact the ATCC
tel#703-365-2700 for further information
Seq primer: M13 Reverse.
Location/Qualifiers
1. .531
/organism="Rattus sp."
/db_xref="taxon:10118"
/clone="RG1CM53"
/clone_lib="Normalized rat embryo, Bento Soares"
/dev_stages="embryo 8, 12, 18 dpc"
/note="Vector: pT7T3Pac; Site_1: EcoRI; Site_2: NotI"
BASE COUNT 133 a 157 c 131 g 110 t
ORIGIN
Query Match 59.2%; Score 472.6; DB 9; Length 531;
Best Local Similarity 93.5%; Pred. No. 1.1e-126;
Matches 493; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
Qy 241 gcaacattacttcgatgcagccatgaacccctgcggagtggtcaagcagaggtgcag 300
Db 5 GCAAGGTTACTTTCATGTGCAGCCATGAATCAGCTGAAGTGGTCAAGCAGCGGATGCAG 64
Qy 301 atgtcaactcaccataccacccggtgcagactgtgtacggcgagtgtgcgaatacaaa 360
Db 65 ATGTCAACTACCCGTACCACCGGTGCAGACTGTGTCCGGGCAGTGGCCAAATGAA 124
Qy 361 ggggcggggcctttaccgcagctacacacccagctgaccatgaacgttctcttccaa 420
Db 125 GGGCGCGGGCCTTTTACCGCAGCTACACGACCCAGCTGACCATGAATGTTCCCTTCCAA 184
Qy 421 gccattcacttcacactatgaattctcgcaggagcactttaaccccccagagcggtac 480
Db 185 GCCATTCACTTTCATACCTATGAGTTCTCTGCAGAGACACTTTTAATCCCGCAGAGCGGTAC 244
Qy 481 aacccaagctcccagctctctctggagcttcgcgcagagctgtagctgcccagaccaca 540
Db 245 AACCCACGCTCCCATGTGCTCTGTGGAGCCTGCGCAGGAGCTGTAGTGCCTGCCGCCACC 304
Qy 541 accccactggagctttgcaaaaacactgtctcaacacccaggagctccttgcttgaactca 600
Db 305 ACCCCACGTGGAGCTTTGCAAAACACTGTCTCAACACCCAGGAGTCCCTGGCCTTGAGCTCA 364
Qy 601 cacattacaggacatatcacaggcatgctcagtcgcttcaggagcgggtatatcaagttagt 660
Db 365 AACATTACAGGACACATCACAGGCATGGCTAATGCTTTCAGGACGCTCTATCAAGTAGGC 424
Qy 661 ggggtgacgcctatttcaggagggtgcagccagagtaatttacagatccctccaca 720
Db 425 GGGGTGACTGCTTACTTCGAGGGGTACAGGCCAGAGTAATTTACCATCCCTCCACA 484
Qy 721 gccatcgcatgctgtgtatgagttcttcaataacctaaactaa 767
Db 485 GCCATCGCATGGCTGGGTATGTAATCTTCAAAATACCTAATACACGAA 531
RESULT 10
BF193023
LOCUS BF193023 563 bp mRNA linear EST 02-NOV-2000
DEFINITION 244452 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.
ACCESSION BF193023
VERSION BF193023.1 GI:11076392
KEYWORDS EST.

SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 563)
AUTHORS Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.
and Keele,J.W.
TITLE Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine
JOURNAL Unpublished (2000)
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR PRIMERS
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCTCCAGTCACGAGC
Plate: 71 row: A column: 23
Seq primer: ATTTAGGTGCACACTATAG.
Location/Qualifiers
1. .563
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 2P1G"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled tissue from testis, ovary,
endometrium, hypothalamus, pituitary, and placenta."
BASE COUNT 140 a 174 c 146 g 103 t
ORIGIN
Query Match 59.0%; Score 470.6; DB 10; Length 563;
Best Local Similarity 95.3%; Pred. No. 4.2e-126;
Matches 485; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
Qy 290 agaggtacagatgcagatcacactcaccatcacccgggtgcagactgtgtacggcgagtg 349
Db 1 AGAGGATCGAGTGAACAACATCACCTACCACCGGGTGACAGACTGTGTACGGGCGAGTGT 60
Qy 350 ggcataaatgaaggggcgccggccttttaccgcagctacacacccagctgaccatgaacg 409
Db 61 GGCAAAAGCAAGGGCGCGGGCCTTTTACCGCAGCTACACCACCCAGCTCAACATGAACG 120
Qy 410 ttcccttccaagcattcacttcacgtacgtatgaattcctcagagcacttaaccccc 469
Db 121 TTCCTCTCCAGCCATCCACTTCATGACCTATGAATTCCTGCAGGAGCACTTTAACCCCC 180
Qy 470 agagacgtgtacaaacccacagctcccacgtctctctctgagcttgcgcagagctgagctg 529
Db 181 AGAGACGTGTACNACCCACAGCTCCACGTCCTCTCCGAGAGCTGCCGACGAGCTGTAGCTG 240
Qy 530 ccgcagcacacacccacactggacgtttgtcaaacactgctcaacacccacagagtgctcttgg 589
Db 241 CGCGCGCGCAACCCCACTGGAGGTTTGCATAAACACTGCTCAATACCCAGGAGTCCCTGG 300
Qy 590 ctttgaactcaacatttacaggacatatacaaggcatggctagtccttcaggagcggtat 649
Db 301 CTTTGAACTCAAACATTACAGGACACATACAGGCATGGCTAGTGCCTTCAGGACGGTGT 360
Qy 650 atcaagtaggtgggtgacccctatttccagggttcacaggggtcagccagagtaattaccaga 709
Db 361 ATCAAGTAGGGGAGTGAACCGCTTACTTCCGAGGGGTGCAGGCGAGAGTAATTTTACCAGA 420
Qy 710 tcccttcacacgccatcgcatggtctgtgtatgagttcttcaaatcacctaataaa 769
|||||

Db 421 TCCCTCCACGCCATCGGTGTGTGTATGAGTTCTTCAATAACCTGATCACTAAAC 480
 Qy 770 gcaagaagtgaggctgcaagtga 798
 Db 481 GCAAGAAGATGGAGGCGCAGCACTGA 509

RESULT 11
 AW341177
 LOCUS
 DEFINITION x25e05.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:2871968 3',
 similar to TR:Q94638 Q94638 MITOCHONDRIAL SOLUTE CARRIER. ;, mRNA
 sequence.
 ACCESSION AW341177
 VERSION AW341177.1 GI:6837803
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 581)
 REFERENCE NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 TITLE Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html

Possible reversed clone: similarity on wrong strand
 Seq primer: -40UP from Gibco
 High quality sequence stop: 369.

FEATURES
 source
 1..581
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2871968"
 /clone_lib="NCI_CGAP_Lu24"
 /tissue_type="carcinoid"
 /lab_host="DH10B"
 /note="Organ: lung; Vector: p7T3D-Pac (Pharmacia) with a
 modified polylinker; Plasmid DNA from the normalized
 library NCI_CGAP_Lu5 was prepared, and ss circles were
 made in vitro. Following HAP purification, this DNA was
 used as tracer in a subtractive hybridization reaction.
 The driver was PCR-amplified cDNAs from a pool of 5,000
 clones made from the same library (clones
 1414920-1417991 and 1520904-1522439). Subtraction by Bento
 Soares and M. Fatima Bonaldo."
 BASE COUNT 138 a 172 c 135 g 134 t 2 others
 ORIGIN

Query Match 58.6%; Score 468; DB 9; Length 581;
 Best Local Similarity 97.5%; Pred. No. 2.4e-125;
 Matches 474; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 279 agtggtcaagcagagtgatgtacaactaccatcacccgggtgacagactgtgt 338
 Db 95 AGTGTCAAGCAGAGAGTGCAGAGTGTACAACCTACCATACCACCGGGTGACAGACTGTGT 154
 Qy 339 acgggagctgtggcaaaatgaaggccggggccctttaccgcagctacacacccagct 398
 Db 155 ACGGGCAGCTGTGGCAAAATGAAGGGCGGGGCCCTTTACCGCAGCTACACCCAGCT 214
 Qy 399 gaccatgaacgttctcttcccaagccattcacttcatgacctatgaattcctgcaggagca 458

Db 215 GACCATGAACGTTCTCTTCCAAAGCCATTCACTTCATGACCTATGAATTCCTGCAGGAGCA 274
 Qy 459 ctttaacccccagagagcgggtacaacccaagctcccaagctctctctgtgagcttgcgagg 518
 Db 275 CTTTAAACCCCCAGAGAGCGGTACAACCCAAAGCTCCACAGTCTCTCTGGAGTTCCGCGAGG 334
 Qy 519 agctgtagctgcgcgacgccacaacccccactgtgacgtttgcaaaacactgctcaacaccca 578
 Db 335 ACCTGTAGCTNGCCGACCCACCAACCCCACTGGACGTTTGCAAAACACTGCTCAACACCCA 394
 Qy 579 ggaagctcttggtttgaactcacacattacaggacatatcacaggcatggtgtagtgcctt 638
 Db 395 GGAGTCTCTGGCTTTGAACTCACACATTACAGAGACATATCACAGGCATGGCTAGTGCTT 454
 Qy 639 caggacggtatatcaagtagtggtgggtgacgcctatttcgaggggtgcagggccagagt 698
 Db 455 CAGGACCGTATATCAAGTAGGTGGGTGACGCCCTATTTTCGAAGGGTGCAGGCCAGAGT 514
 Qy 699 aattaccagatccctccacagccatcgcatggtctgtgtatgagttcttcaataacct 758
 Db 515 AAATTACAGATCCCTCCACAGCCATCGCATGGTCTGTATGAGTTCTCCANATACCT 574
 Qy 759 aatcac 764
 Db 575 AATCAC 580

RESULT 12
 BF143133
 LOCUS
 DEFINITION 601788025F1 NCI_CGAP_Lu30 Mus musculus cDNA clone IMAGE:4015758 5',
 mRNA sequence.
 ACCESSION BF143133
 VERSION BF143133.1 GI:10982173
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 830)
 REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM9262 row: a column: 07
 High quality sequence stop: 607.

FEATURES
 source
 1..830
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:4015758"
 /clone_lib="NCI_CGAP_Lu30"
 /tissue_type="tumor, metastatic to mammary"
 /lab_host="DH10B"
 /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; transgenic model WNT-1, expression driven by
 MMTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo
 dt. Library constructed by Life Technologies.
 Investigator providing samples: Gilbert Smith, NIH"
 BASE COUNT 202 a 239 c 220 g 169 t
 ORIGIN

Query Match 56.8%; Score 453.6; DB 10; Length 830;
 Best Local Similarity 92.4%; Pred. No. 4.4e-121;
 Matches 477; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 283 gtcaagcagagatcagatgtacaaactcacacaccacccgggtgacagactgtgtacgg 342
 Db 1 gtcaagcagagatcagatgtacaaactcacacaccacccgggtgacagactgtgtcgg 60

QY 343 gcaagtgtggcaaaatgaaggccggggccttttaccgcagctacacaccagctgacc 402
 Db 61 gcagtggtggcaaaatgaaggccggggccttttaccgcagctacacaccagctgacc 120

QY 403 atgaacgttccttccaaagccattcaactatcagactatgaattcctcagagacacttt 462
 Db 121 atgaatgtcccccttccaaagccattcaactatcagactatgaattcctcagagacacttt 180

QY 463 aacccccagagcgtacaaacccagctcccaagctccctctgtgagctgcagagact 522
 Db 181 aacccccagagcgtacaaacccagctcccaagctccctctgtgagctgcagagact 240

QY 523 gtactgtgcgcagcagcaacccaccactggacgttttgcaaaacactgtcacaacccagag 582
 Db 241 gtactgtgcgcagcagcaacccaccactggacgttttgcaaaacactgtcacaacccagag 300

QY 583 tccctgtgttgaaactcacacattacagagacatatcagagcagtgcttagtcctcagg 642
 Db 301 tccctgtgttgaaactcacacattacagagacatatcagagcagtgcttagtcctcagg 360

QY 643 acggtatatcaagttagtgagggtgacccctatttccagagggtgcagccagagtaatt 702
 Db 361 acggttatcaagttagtgagggtgacccctatttccagagggtgcagccagagtaatt 420

QY 703 taccagatccccctcacagcagcagctgctgtgtgtatgagttctcaaatcacctaatt 762
 Db 421 taccagatccccctcacagcagcagctgctgtgtgtatgagttctcaaatcacctaatt 480

QY 763 actaaaggcgaagaagtagtgagggtgcgaagtga 798
 Db 481 actaaaggcgaagaagtagtgagggtgcgaagtga 516

RESULT 13
 BM483098
 LOCUS 536233 MARC 3BOV Bos taurus cdna 5', mRNA linear EST 05-FEB-2002
 DEFINITION
 ACCESSION BM483098
 VERSION BM483098.1 GI:18533417
 KEYWORDS EST.
 SOURCE cow.
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.
 1 (bases 1 to 578)
 Smith, T.P.L., Grosche, W.M., Freking, B.A., Roberts, A.J., Stone, R.T.,
 Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C., Bennett,
 G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A., Chitko-McKown, C.G.,
 Perte, G., Holt, I., Karamycheva, S., Liang, F., Quackenbush, J. and
 Keele, J.W.
 Sequence evaluation of four pooled-tissue normalized bovine cdna
 libraries and construction of a gene index for cattle
 Genome Res. 11 (4), 626-630 (2001)
 21180013
 Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smithemail.marc.usda.gov
 Single pass sequencing. Bases called and alt_trimmed with phred
 v0.980904.e. Vector identified by cross_match with the -minscore 18
 and -minmatch 12 options.
 PCR Primers

FORWARD: AGAAACAGCTATGACCAT
 BACKWARD: GTTTCCTCCAGTACGACG
 Plate: 5 row: F column: 16
 Seq primer: ATTTAGGTGACACTATAG.
 Location/Qualifiers
 1..578
 /organism="Bos taurus"
 /db_xref="taxon:9913"
 /clone_lib="MARC 3BOV"
 /tissue_type="pooled"
 /lab_host="DHI0B"
 /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
 Library made from pooled tissue from marrow, alveolar
 macrophage, ovary, fetal semitendinosus muscle, and fetal
 longissimus muscle."
 139 a 184 c 141 g 114 t

BASE COUNT
 ORIGIN

Query Match 56.8%; Score 453.4; DB 10; Length 578;
 Best Local Similarity 94.7%; Pred. No. 4.4e-121;
 Matches 469; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 304 tacaactcacatcacaccgggtgacagactgtgtacgggagtggtgcaaaatgaagg 363
 Db 1 TACAACCTCACCGTACCACCGGTGACAGACTGTGTCCGGGAGTGTGGCAGAACGAAGG 60

QY 364 gccggggcctttaccgcagctacacacccagctgacacgaacgttccttccaaagcc 423
 Db 61 gctggggccttttaccgcagctacacacccagctgacacgaacgttccttccaaagcc 120

QY 424 attcactcagacatgaattcctgcagagcacttaaccccgagacggtacacac 483
 Db 121 ATTCACTTCATGACCTACGAATTCCTGCAGGAGCACTTTAACCCCGAGAGCGGTACAC 180

QY 484 ccaagctccccagctcctctctgtggagctgtgcagagagctgtgctgcgcagcgacacac 543
 Db 181 CCCAGCTCCACGCTCTCTCCGGAGCCTGTGCAGGAGCTGTAGTGTCCGCTGCCACAACC 240

QY 544 ccactggagcttgcacaaacactgctcaacccagggagctcctgtgttgaactcacac 603
 Db 241 CCCTGGAGCTTGCACAAACACTGCTCAACCCAGGAAATCCCTGGCTTTGAACCTCAAC 300

QY 604 attacagacatatcacaggcatggctgctcctcagagcgttatatcaagtaggtgg 663
 Db 301 CTTACAGGACATATCACAGGCATGGCTAGTGTTCAGGAGCGGTGTATCAAGTAGCGCGG 360

QY 664 gtgaccgcctatttcgaggggtgcagggcagagtaattaccagatccccctccacagcc 723
 Db 361 GTGACTGCCTACTTCCGAGGGGTGCAGGCTAGAGTCAITTTACCAGATCCCTCCACAGCC 420

QY 724 atcgcatgctgtgtatgagttcttcaaaacacataacataacataaaggcaagagagtg 783
 Db 421 ATCGCTGGCTGTGTATGAGTTCTTCAAAATACCTGTACTCAAAAGGCAAGAGAGTGG 480

QY 784 agggctggcaagtga 798
 Db 481 AGGCAGGCAAGTGA 495

RESULT 14
 AL530804
 LOCUS AL530804 LTI_NFL001_NBC4 Homo sapiens cdna clone CS0DD008YH17 5
 DEFINITION prime, mRNA sequence.
 ACCESSION AL530804
 VERSION AL530804.1 GI:12794297
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 888)

AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1. .888
/location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CSDD008YH17"
/clone_lib="LTI_NFL001_NBC4"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 168 a 244 c 322 g 139 t 15 others
ORIGIN

Query Match 54.8%; Score 437.2; DB 9; Length 888;
Best Local Similarity 97.2%; Pred. No. 2.7e-116;
Matches 459; Conservative 6; Mismatches 5; Indels 2; Gaps 2;

Qy 1 atcagaagtacagctaccagctgcccctctatgcgaatgtgttgaggccctctg 60
Db 417 ATGCAGAGCTACAGCTGACCCAGCTGCCGCTATCGCAATGTTGGAGGCCCTCTG 476

Qy 61 aggattataagaacggggcctatgagggccatgaggggctgaaacgtcacagcaaca 120
Db 477 AGGATTATAAGAAGCGGCTATGAGGCCCATGAGGGGCTGAACGTCACAGCAACA 536

Qy 121 ggcgagggcctgccacgccttatttgctgtctacgaaagttaaaagacattg 180
Db 537 GCGCAGGGCTGCCACCGCCCTTATTTGCTGTCBACGAAAGTTAAAAAGACATG 596

Qy 181 agtatgaatacaccctggggcaatagccattatgccaatggtgcggcggtgtgtg 240
Db 597 AGTATGTTATCCACCTGGGGCAATAGCCATATTCCCAATGTCGGCGGCTGTG 656

Qy 241 gcaacattacttcgatgcagccatgaacccctgcggaagtgttcaagcagaggtacg 300
Db 657 GCAACATTACTTCATGATGAGCCATGACCCCTGCGGAAGTGTGTCAGSAGAGGATCGAG 716

Qy 301 atgtacaactcaccataccaccgggtgacagactgtgtacggcgagtggt-ggcaaaatga 359
Db 717 ATGTACAACCTACCACACCGGGTGACAGACTGTGTACGGCGAGTGTGGCAAAATGA 776

Qy 360 agggcgggggccttttacgcagcactacaccacccactgaccatgaactctctcca 419
Db 777 AGGGGGCGGGGCCCTTTTACCGCAGCTACACACCCCAATGACGACGTCCTCTTCCA 836

Qy 420 agccattacctca-tgacctatgaattcctgcaggagcactttaaccacca 470
Db 837 AGCCATTACTTCAWTGACCTATGAATCTCTGCAGGAGCCCTTTTAAACCCCA 888

RESULT 15
BF515180 451 bp mRNA linear EST 07-DEC-2000
LOCUS UI-H-BW1-anl-b-10-0-UI.sl NCI_CGAP_Sub7 Homo sapiens cDNA clone
DEFINITION IMAGE:3082698 3', mRNA sequence.
ACCESSION BF515180

BF515180.1 GI:11600447
EST.
human.
Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 451)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. cDNA Library Preparation: M.B. Soares Lab Clone distribution:
NCI-CGAP clone distribution information can be found through the
I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbrrp/image/image.html
Seq primer: M3 Forward
POLYA=Yes.

Location/Qualifiers
1. .451
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:3082698"
/lab_host="NCI_CGAP_Sub7"
/clone_lib="NCI_CGAP_Sub7"
/notes="Vector: pT7T3D-Pac (Pharmacia).with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; NCI_CGAP_Sub7
is a subcloned library derived from NCI_CGAP_Sub6. The
NCI_CGAP_Sub7 library had 12 million recombinants. A
single-stranded DNA preparation of NCI_CGAP_Sub6 was used
as a tracer in a subtractive hybridization with a driver
comprising: the IMAGE pool (NCI_CGAP_Kid3 pool 1 LLAM
3334-3337, 3682-3683, 3798-3803 (IMAGE CloneIDs
132376-132391), 1456008-1456775, 1500552-1502855);
NCI_CGAP_Kid5 pool 1 LLAM 3338-3342, 3722-3725, 3776-3778
(IMAGE CloneIDs 1323912-1325831, 1471368-1472903,
1492104-1493255); NCI_CGAP_Lu5 pool 1 LLAM 3575-3582,
3851-3854 (IMAGE CloneIDs 1414920-1417991, 1520904-1522439
); NCI_CGAP_GC4 pool 1 LLAM 3164-3167, 3716-3720,
3733-3735 (IMAGE CloneIDs 1257096-1258631, 1469064-1470983
, 1475592-1476743); NCI_CGAP_Pr22 pool 1 LLAM 2457-2459,
2758-2759, 3062-3068 (IMAGE CloneIDs 985608-986759,
1101192-1101959, 1217928-1220615); NCI_CGAP_Co10 pool 1
LLAM 2644-2653, 2871-2872 (IMAGE CloneIDs 1057416-1061255
, 1144584-1145351). (6% of the driver population), plus a
pool of 3,840 arrayed clones from NCI_CGAP_Sub1 (IMAGE
CloneIDs 2708616-2710535) and NCI_CGAP_Sub2 (IMAGE
CloneIDs 2710536-2712455) (4% of the driver population
, plus a pool of 11,136 clones from NCI_CGAP_Sub3 (IMAGE
CloneIDs 2712456-2723591) (10% of the driver population),
plus a pool of 5,472 clones from NCI_CGAP_Sub4 (IMAGE
CloneIDs 2723592-2729326) (40% of the driver population),
plus a pool of 4032 clones from NCI_CGAP_Sub6 (IMAGE
CloneIDs 2728969-2733190) (40% of the driver population).
Subtraction was performed as previously described (Bonaldo
Lennon & Soares (1996): Normalization and Subtraction:
Two Approaches To Facilitate Gene Discovery. Genome
Research 6, 791-806.
TAG_LIB=NCI_CGAP-Lu13
TAG_TISSUE=lung
TAG_SEQ=GCCGG"

BASE COUNT 112 a 129 c 109 g 101 t
ORIGIN

Query Match 54.5%; Score 435; DB 10; Length 451;
Best Local Similarity 100.0%; Pred. No. 9e-116;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 223 ggtgcggccgggtgtgtgtggaacattacttcatgatgcagccatgaaccctgcggaagtg 282
Db 17 GGTGCGCGCGGTGTGTGTGGCAACATTTACTTCATGTGCAGCCATGAACCTTGCAGAGTG 76
Qy 283 gtcagagagagatgcagatgtacaactcaccataccaccgggtgacagactgtgtacgg 342
Db 77 GTCAGCAGAGGATGCAGATGTACAACCTCACCATACCACCGGGTGACAGACTGTGTACGG 136
Qy 343 gcagtgtggcaaaatgaaggggccggggcccttttacgcgcagctacaccaccagctgacc 402
Db 137 GCAGTGTGGCAAAATGAAGGGGCGGGGCCCTTTACCGCAGCTACACCACCAGCTGACC 196
Qy 403 atgaacgttcctttcccaagccattcacttcatgacctatgaattcctgcaggagcacttt 462
Db 197 ATGAACGTTCCCTTTCCAAAGCCATTTCACCTTCATGACCTATGAATTCCTCAGGAGCATT 256
Qy 463 aacccccagagacggtacaaaccccaagctcccaagctccctctctgtgagcttgccaggagct 522
Db 257 AACCCCCAGAGACGGTACAAACCAAGCTCCACACCTCTCTCTGGAGCTTGGCGCAGGAGCT 316
Qy 523 gtacgtgcgcagccacaacccccactgacgttttgcaaaacactgctcaacacccaggag 582
Db 317 GTAGCTGCGGAGCCACAAACCCGACTGGACGTTTGCAAAACACTGCTCAACACCCAGGAG 376
Qy 583 tccttggtttgaactcaacacattacaggacatatcacaggcatggctagtgccttcagg 642
Db 377 TCCTTGGCTTTGAACTCACACATTACAGGACATATCACAGGCATGGCTAGTGCCTTCAGG 436
Qy 643 acgtatatcaagta 657
Db 437 ACGGTATATCAAGTA 451

Search completed: September 28, 2002, 06:22:48
Job time: 9189 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2002, 07:57:11 ; Search time 5287.54 Seconds
(without alignments)
1127.947 Million cell updates/sec

Title: US-09-870-113-9
Perfect score: 285
Sequence: 1 atgcagagtctacagctga.....tgaacctgcggaaggtga 285

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.ot.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg_hum.*
31: em.htg_inv.*
32: em.htg_other.*
33: em.htgo_inv.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
------------	-------	-------	-------	--------	----	-------------

ALIGNMENTS

RESULT	1					
AX061229	LOCUS	AX061229	Sequence 76 from Patent WO0078953.	1322 bp	DNA	linear PAT 22-JAN-2001
AX061229	DEFINITION	Sequence 76 from Patent WO0078953.				
AX061229	ACCESSION	AX061229				
AX061229.1	VERSION	AX061229.1	GI:12406365			
KEYWORDS						
SOURCE		human.				
ORGANISM		Homo sapiens				
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS		1 (bases 1 to 1322) Lal, P., Yang, J., Yue, H., Hillman, J.L., Tang, Y.T., Bandman, O., Burford, N., Baughn, M.R., Azimzai, Y., Lu, D.A., Au-Young, J. and Patterson, C.				
TITLE		Human transport proteins				
JOURNAL		Patent: WO 0078953-A 76 28-DEC-2000;				
FEATURES		Incyte Genomics, Inc. (US)				
source		Location/Qualifiers				
		1..1322				
		/organism="Homo sapiens"				
		/db_xref="taxon:9606"				
BASE COUNT		328 a	376 c	351 g	267 t	

FT CDS 30..1124
/gene="HMR3/4"
/product="mitochondrial RNA splicing protein 3/4"
/function="putative mitochondrial solute carrier"
/protein_id="CAC27996.1"
/translation="MELEGRGAGVAGGPAAGPGRSPGESALLDGNLQVGRGAGGGE
AGACRPVPRQDPDGYEALPAGATYTHMVAGAVAGILEHCYVMPIDCVKTRQMSLQ
PDPAARNVLEALMRIIRTEGLRNPRLNVTATGAPAHALYFACYEKLKLTLSQVI
HPGNSHLANGAACCVATLLHDAAMNPAEVVKQRMQVNSPYHRVTCVRAVWONEGAG
AFYRSYVLTQMNVPFOAIHMTVEFLQEHENPORRYNPSHVLGACAGAVAAATTP
LDVCKTLNTOESLALNSHITGHTGASAFRTYQVGGVTAYFRGVQARVIYQIPSTA
IANSVYEFFKLIITRQEEWRAGK"

SQ Sequence 1448 BP; 322 A; 408 C; 435 G; 283 T; 0 other;

Query Match 98.2%; Score 280; DB 17; Length 1448;
Best Local Similarity 100.0%; Pred. No. 6.6e-83;
Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgcagagtctacagcctgaccagctgcccgcctatcgcaatgtgttgaggccctctgg 60
|||||
DB 327 ATGCAGAGTCTACAGCTGACCGACGCTGCCGCTATCGCAATGTGTGGAGGCCCTCTGG 386

QY 61 aggattataagaacggagggccctatgaggggccctatgaggggctgaacgtcacagcaaca 120
|||||
DB 387 AGGATTATAAGAACGAGGAGGCGCTATGGAGGCCCATGAGGGGCTGAACGTCACAGCAACA 446

QY 121 ggcgcagggcctgcccagccctttatcttgcctactacgaaagttaaaaagacattg 180
|||||
DB 447 GCGCAGGGCCTGCGCCAGCCCTTTATTTGCTCTGCTAGAAAAGTTAAAAAGACATTG 506

QY 181 agtgatgaataccaccctggggccaatgccaattggtcgccgggtgtgtg 240
|||||
DB 507 AGTGATGTAATCCACCCTGGGGCAATAGCCATATTGCCAATGTGTGGCCGGGTGTGTG 566

QY 241 gcaacattactcatgctgagccatgaacctgcggaag 280
|||||
DB 567 GCAACATTACTTCAATGTCAGCCCATGAACCTCGCGGAAG 606

RESULT 4
AX071440 401 bp DNA linear PAT 25-JAN-2001
LOCUS
DEFINITION Sequence 1912 from Patent WO0102568.
ACCESSION AX071440
VERSION AX071440.1 GI:12581791

KEYWORDS human.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 401)
AUTHORS Williams,L.T., Escobedo,J., Innis,M.A., Garcia,P.D., Klinger,J.,
Kassam,A., Reinhard,C., Randazzo,F., Kennedy,G.C., Pot,D.,
Lanson,G., Reinhar,R., Crkenjakov,R., Drmanac,S., Dickson,M.,
Labat,I., Leshkowitz,D., Kita,D., Garcia,V. and Strache-Crain,B.
TITLE Human genes and gene expression products
JOURNAL Patent: WO 0102568-A 1912 11-JAN-2001;
CHIRON CORPORATION (US); HYSEQ, INC. (US)

FEATURES
source
1..401
/organism="Homo sapiens"
/db_xref="taxon:9606"
misc_feature 1..401
/note="n = A,T,C or G"
BASE COUNT 100 a 104 c 120 g 76 t 1 others
ORIGIN

Query Match 95.8%; Score 273; DB 6; Length 401;
Best Local Similarity 100.0%; Pred. No. 1.2e-80;
Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 gtctacagctgacccagctgccgctatcgcaatgtgttgaggccctctgagagatta 67
|||||
DB 9 GTCACAGCTGACCCAGCTGCCGCTATCGCAATGTGTGGAGGCCCTCTGGAGATTA 68
|||||
QY 68 taagaacggagggcctatgagggcccatgaggggctgaacgtcacagcaacagcgcaag 127
|||||
DB 69 TAAGAAGGAGGGCCTATGAGGGCCCATGAGGGGCTGAACGTCACAGCAACAGGCGCAG 128
|||||
QY 128 ggcctgcccagcctttatcttgcctactacgaaagttaaaaagacattgagtgatg 187
|||||
DB 129 GGCCTGCCACGCCCTTTATTTGCTCTGCTAGAAAAGTTAAAAAGACATTGAGTGATG 188
|||||
QY 188 taatcacccctggggccaatgccaattgccaatggtgcgcccgggtgtgtgccaacat 247
|||||
DB 189 TAATCACCCCTGGGGCAATAGCCATATTGCCAATGTGTGGCCGGGTGTGTGGCAACAT 248
|||||
QY 248 tactctatgagcagcccatgaacctgcggaag 280
|||||
DB 249 TACTTCATGATGACGCCCATGAACCCCTGCGGAAG 281

RESULT 5

AF327403 1889 bp mRNA linear PRI 02-MAY-2001
LOCUS
DEFINITION Homo sapiens putative mitochondrial solute carrier splice variant
mRNA, complete cds, alternatively spliced, nuclear gene for
mitochondrial product.

ACCESSION AF327403
VERSION AF327403.1 GI:13926049

KEYWORDS
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1889)

AUTHORS Li,F.Y., Nikali,K., Gregan,J., Leibiger,I., Leibiger,B.,
Schweyen,R., Larsson,C. and Suomalainen,A.
TITLE Characterization of a novel human putative mitochondrial
transporter homologous to the yeast mitochondrial RNA splicing
proteins 3 and 4

JOURNAL FEBS Lett. 494 (1-2), 79-84 (2001)
MEDLINE 21195335
PubMed 11297739

REFERENCE 2 (bases 1 to 1889)

AUTHORS Li,F., Nikali,K., Gregan,J., Leibiger,I., Leibiger,B., Schweyen,R.,
Larsson,C. and Suomalainen,A.
Direct Submission

TITLE Human Molecular Genetics, National Public
JOURNAL Submitted (08-DEC-2000) Human Molecular Genetics, National Public
Health Institute, Mennerheimintie 166, Helsinki 00300, Finland

FEATURES
Location/Qualifiers

source
1..1889

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="10"

/map="10q24"

1032..1565

CDS
/note="similar to mitochondrial RNA splicing protein 3/4;
alternatively spliced"

/codon_start=1

/product="putative mitochondrial solute carrier splice
variant"

/protein_id="AAK49520.1"

/db_xref="GI:13926050"

/translation="MNPAEVVKQRMQVNSPYHRVTCVRAVWONEGAGAFYRSYTTQ
LTMNVPQAIHFMTYQLQEHFNPNRSHVLSGACAGAVAAATPLDVCCKTL
NTOESLALNSHITGHTGASAFRTYQVGGVTAYFRGVQARVIYQIPSTAIANSVYE
FFKYLITRQEEWRAGK"

BASE COUNT 404 a 508 c 544 g 433 t
ORIGIN

Query Match 78.9%; Score 225; DB 9; Length 1889;

```
Best Local Similarity 100.0%; Pred. No. 2.2e-64;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atgcagagctctacagcctgaccagctgcccgcctatgcgaatgtgttgagccctctgg 60
Db 306 ATGCAGAGCTCTACAGCCTGACCCAGCTGCCGCTATGCAATGTGTGGAGGCCCTCTGG 365
Qy 61 aggtattataagaacgagggcctatggagggccatgaggggctgaacgtcacagcaaca 120
Db 366 AGGATTATTAAGACGAGGGCCTATGGAGGCCCATGAGGGGCTGAAGCTCACACACA 425
Qy 121 ggcgagggcctgcccagcccttattttgctgctacgaaaaagttaaaaaagacattg 180
Db 426 GCGCAGGGCCTGCCAGCCCTTTATTTGCTGCTACGAAAAGTTAAAAAGACATTG 485
Qy 181 agtgaatgaatccaccctgggggcaatagccatattgccaatggt 225
Db 486 AGTGATGTAATCCACCTGGGGCAATAGCCATATTGCCAATGGT 530

RESULT 6
HSA303078
ID HSA303078 standard: RNA; HUM; 1889 BP.
XX AC
XX AJ303078;
SV AJ303078.1
XX 02-FEB-2001 (Rel. 66, Created)
DT 26-APR-2001 (Rel. 67, Last updated, Version 2)
XX Homo sapiens mRNA for mitochondrial RNA splicing protein 3/4 (HMRS3/4
DE gene), 1889 bp cDNA splice variant
XX HMRS3/4 gene; mitochondrial RNA splicing protein 3/4.
XX Homo sapiens (human)
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
XX [1]
RN 1-1889
RA Nikali K.;
RT Submitted (12-DEC-2000) to the EMBL/GenBank/DBJ databases.
RL Nikali K., Human Molecular Genetics, National Public Health Institute,
RL Mannerheimintie 166, Helsinki, FI-00300, FINLAND.
XX [3]
RN Li F., Nikali K., Gregan J., Leibiger I., Leibiger B., Schweyen R.,
RA Larsson C., Suomalainen A.;
RT "Characterization of a novel human putative mitochondrial transporter
RT homologous to the yeast mitochondrial RNA splicing proteins 3 and 4";
RL FEBS Lett. 494:79-84(2001).
XX
FH Key Location/Qualifiers
FH 1. .1889
FT /db_xref="taxon:9606"
FT /organism="Homo sapiens"
FT CDS 1032..1565
FT /gene="HMRS3/4"
FT /product="mitochondrial RNA splicing protein 3/4"
FT /function="putative mitochondrial solute carrier"
FT /protein_id="CAC27997.1"
FT /translation="MNPAEVVKRMQMNYSPIHRVDCVRAVQWNEGAFYRSYITQL
FT TMNVPFQAIHFMYEFLQEHFNFORRNPSSHLVSGACAVAAAATPLDVCKTLNT
FT QESLALNSHTIGTITGMASAFRTYVQGVGTAYFRGVQARVIYQIPSTAIWASYIEFFK
FT YLITKROEWRAGK"
XX
SQ Sequence 1889 BP; 404 A; 508 C; 544 G; 433 T; 0 other;
```

```
Query Match 78.9%; Score 225; DB 17; Length 1889;
Best Local Similarity 100.0%; Pred. No. 2.2e-64;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atgcagagctctacagcctgaccagctgcccgcctatgcgaatgtgttgagccctctgg 60
Db 306 ATGCAGAGCTCTACAGCCTGACCCAGCTGCCGCTATGCAATGTGTGGAGGCCCTCTGG 365
Qy 61 aggtattataagaacgagggcctatggagggccatgaggggctgaacgtcacagcaaca 120
Db 366 AGGATTATTAAGACGAGGGCCTATGGAGGCCCATGAGGGGCTGAAGCTCACACACA 425
Qy 121 ggcgagggcctgcccagcccttattttgctgctacgaaaaagttaaaaaagacattg 180
Db 426 GCGCAGGGCCTGCCAGCCCTTTATTTGCTGCTACGAAAAGTTAAAAAGACATTG 485
Qy 181 agtgaatgaatccaccctgggggcaatagccatattgccaatggt 225
Db 486 AGTGATGTAATCCACCTGGGGCAATAGCCATATTGCCAATGGT 530

RESULT 7
AL353719/c
LOCUS
DEFINITION Human DNA sequence from clone RP11-85A1 on chromosome 10, complete
sequence.
ACCESSION AL353719 AC007643
VERSION AL353719.10 GI:15787725
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 123160)
AUTHORS Ramsay, H.
TITLE Direct Submission
JOURNAL Submitted (25-SEP-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
COMMENT On Sep 26, 2001 this sequence version replaced gi:14280413.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em: EMBL; Sw:
SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 10, constructed by the Sanger Centre Chromosome 10
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr10
RP11-85A1 is from the library RPC1-11.1 constructed by the group of
Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6
IMPORTANT: This sequence is not the entire insert of clone
RP11-85A1. It may be shorter because we sequence overlapping
sections only once, except for a short overlap.
The true left end of clone RP11-85A1 is at 1 in this sequence. The
true left end of clone RP11-483F1 is at 123061 in this sequence.
The true right end of clone RP11-129J12 is at 51589 in this
sequence.
```

FEATURES	Location/Qualifiers									
source	1. .123160									
misc_feature	/organism="Homo sapiens"									
	/db_xref="taxon:9606"									
	/chromosome="10"									
	/clone="RP11-85A1"									
	/clone_lib="RPC1-11.1"									
unsure	complement(7086. .7118)									
	/note="Single clone region. Sequence from reads from a short insert library derived from a single pUC clone. Restriction digest data confirm the assembly."									
	52162									
	52167									
	BASE COUNT 33221 a 28084 c 27196 g 34659 t									
ORIGIN										
Query Match 78.9%; Score 225; DB 9; Length 123160;										
Best Local Similarity 100.0%; Pred. No. 4.2e-64;										
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;										
QY	1	atcagagctcacagctgaccagctgcccgcctatcgcaatgtgttgaggccctctgg 60								
Db	85540	ATGCAGAGTCTACAGCTGACCCAGCTGCCGCTATCGCAATGTGTGAGGCCCTCTGG 85481								
QY	61	aggattataaagacgagggccctatggaggcccatgaggggctgaacgctcacgcaaca 120								
Db	85480	AGGATTATAAAGAACGGAGGCCATATGGAGGCCCATGAGGGGCTGAAGCTCACGCAACA 85421								
QY	121	ggcgagggcctgccagccctttatttgcctctctacgaaagttaaaaagacattg 180								
Db	85420	GGCGAGGGCCTGCGCCACGCGCCCTTTATTTCGCTGCTACGAAAGTTAAAAAGACATTG 85361								
QY	181	agtgatgaatccaccctgggggccaatgccaattgccaatggt 225								
Db	85360	AGTGATGTAATCCACCTGGGGGCATAGCCATATGTCATGCTGTCATGCT 85316								
RESULT	8									
LOCUS	AC096351/c									
DEFINITION	Rattus norvegicus chromosome Rf1 clone CH230-24M6, WORKING DRAFT									
ACCESSION	AC096351									
VERSION	AC096351.2 GI:17944054									
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.									
SOURCE	Norway rat.									
ORGANISM	Rattus norvegicus									
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.									
AUTHORS	1 (bases 1 to 221062)									
Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C., Albrooks, S.L., Amarantunge, H.C., Are, J.R., Banks, T., Barbara, J., Benton, J., Blinag, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowle, S., Brlewa, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escott, M., Falls, T., Ferraguto, D., Flagagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kresh, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Lichtarge, O., Lieu, C., Liu, J., Liu, W.,										

Loulseged, H., Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogulu, M., Okwuonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sisson, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G. and Gibbs, R.

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GE2D
Center clone name: CH230-24M6
----- Summary Statistics
Assembly program: Phrap; version 0.990329First call to findPhrapList
Consensus quality: 200164 bases at least Q40
Consensus quality: 205303 bases at least Q30
Consensus quality: 209366 bases at least Q20
Estimated insert size: 209880; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-1p estimation
Quality coverage: 3.9x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 33 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
1
32851: contig of 32851 bp in length
32951: gap of unknown length
32952 32951: contig of 16033 bp in length
32952 48984: gap of unknown length
48985 49084: gap of unknown length
49085 63815: contig of 14731 bp in length
63815 63915: gap of unknown length
63916 76005: contig of 12090 bp in length
76006 76105: gap of unknown length
76106 88841: contig of 12736 bp in length
88842 88941: gap of unknown length
88942 100591: contig of 11650 bp in length
100592 100691: gap of unknown length
100692 111702: contig of 11011 bp in length
111703 111802: gap of unknown length
111803 122665: contig of 10862 bp in length
122665 122764: gap of unknown length
122765 129711: contig of 6947 bp in length
129712 129711: gap of unknown length

[illegible]

RESULT 13
AC108878

```
* 182589 182608: gap of unknown length
* 182609 183975: contig of 1367 bp in length
* 183976 183995: gap of unknown length
* 183996 186393: contig of 2398 bp in length
* 186394 186413: gap of unknown length
* 186414 188416: contig of 2003 bp in length
* 188417 188436: gap of unknown length
* 188437 189584: contig of 1148 bp in length
* 189585 189605: gap of unknown length
* 189606 190648: contig of 1044 bp in length
* 190649 190668: gap of unknown length
* 190669 191702: contig of 1034 bp in length.
FEATURES
    Location/Qualifiers
        1..191702
            /organism="Mus musculus"
            /strain="C57BL6/J"
            /db_xref="taxon:10090"
            /clone="RP23-256H24"
            /sex="male"
BASE COUNT 53763 a 43806 c 42665 g 51108 t 360 others
ORIGIN
Query Match 38.5%; Score 109.8; DB 2; Length 191702;
Best Local Similarity 68.0%; Pred. No. 2.6e-25;
Matches 153; Conservative 0; Mismatches 72; Indels 0; Gaps 0;
QY 1 atgcagagtctacagcctgaccagctccgcgctatcgaatgtgtgagggccctctgg 60
Db 106283 ATGCAGAGTTTAAATCCAGATCCCAAGCCGGTATACAGCATATATGGCGCCCTCAAG 106342
QY 61 aggattataagaacgagggcctatggagggccatgagggcgtaacgctcacagcaaca 120
Db 106343 AGGATCATGCACACTGAAGGCTTCTGGAGGCCCTCGGGGCGCTGAAGCTGATGATG 106402
QY 121 ggcgcagggcctgcacagccctcttatttgcctgctacgaaagttaaagaagacattg 180
Db 106403 GGTGAGGGCCCGCGCAGCCCATGATTTTGGCTGCTATGAACATGAAAGGACTTTA 106462
QY 181 agtgaatgaatccaccctgggggcaatagccatatgtccaatggt 225
Db 106463 AATGACGTTTTTCAGCCACCACAGGAACAGCCCATCTAGCTAATGTT 106507
RESULT 14
AX306258
LOCUS AX306258 483 bp DNA linear PAT 11-DEC-2001
DEFINITION Sequence 1009 from Patent WO0188188.
ACCESSION AX306258
VERSION AX306258.1 GI:17645513
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (sites)
AUTHORS Ishikawa,K., Asai,S., Takahashi,Y., Nagata,T. and Ishii,Y.
TITLE Method for examining ischemic conditions
JOURNAL Patent: WO 0188188-A 1009 22-NOV-2001;
School Juridical Person Nihon University (JP)
FEATURES
    Location/Qualifiers
        1..483
            /organism="Mus musculus"
            /db_xref="taxon:10090"
BASE COUNT 127 a 139 c 121 g 96 t
ORIGIN
Query Match 36.9%; Score 105.2; DB 6; Length 483;
Best Local Similarity 93.2%; Pred. No. 3.5e-24;
Matches 110; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 165 gttaaaaaagacattgagtgtatccaccctgggggcaatagccattgccaattg 224
```

```
Db 15 GTTAAAAAGACATTGAGTGACGTAATCCACCGGGCGCAATAGCCATTATTCGCAATGG 74
QY 225 tgcgcccgggtgtggtgcaacattacttcattcatgacacccatgaacctgggaagcc 282
Db 75 TGCACCGGATGTGTGGGACATTACTTTCATGATGACCCATGAATCCAGCGGAAGTC 132
RESULT 15
AF216674 108765 bp DNA linear HTG 02-JUN-2001
LOCUS Homo sapiens chromosome 8 clone RP1-158P13 map 8p, WORKING DRAFT
DEFINITION SEQUENCE, 16 unordered pieces.
ACCESSION AF216674
VERSION AF216674.3 GI:14280182
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 108765)
AUTHORS Schilhabel,M.B., Baumgart,C., Blechschmidt,K., Dette,M., Jahn,N.,
Lehmann,R., Menzel,U., Polley,A., Reichwald,K., Schudy,A.,
Siddiqui,R., Taudien,S., Wen,G., Schlegelberger,B., Siebert,R.,
Rosenthal,A. and Platzer,M.
TITLE Chromosome 8 genomic sequence
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 108765)
AUTHORS Polley,A., Wen,G., Baumgart,C., Dette,M., Jahn,N., Schilhabel,M.,
Menzel,U. and Rosenthal,A.
JOURNAL Direct Submission
COMMENT Submitted (15-DEC-1999) Genome Analysis, Institute of Molecular
Biotechnology, Reutenbergstrasse 11, Jena 07745, Germany
On Jun 2, 2001 this sequence version replaced gi:8151796.
-----
Center: Institute of Molecular Biotechnology
Center code: IMB
Web site: http://genome.imb-jena.de/
Contact: gscj-submit@genome.imb-jena.de
-----
Project Information
Center project name: H241
Center clone name: RP1-158P13
-----
Summary Statistics
Sequencing vector: M13; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 87772 bases at least Q40
Consensus quality: 95159 bases at least Q30
Consensus quality: 100612 bases at least Q20
Quality coverage: 3.95 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 16 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1101: contig of 1101 bp in length
* 1102 1201: gap of unknown length
* 1202 21139: contig of 19938 bp in length
* 21140 21239: gap of unknown length
* 21240 36201: contig of 14962 bp in length
* 36202 36301: gap of unknown length
* 36302 47101: contig of 10800 bp in length
* 47102 47201: gap of unknown length
* 47202 56806: contig of 9605 bp in length
* 56807 56906: gap of unknown length
* 56907 65083: contig of 8177 bp in length
* 65084 65183: gap of unknown length
* 65184 73521: contig of 8338 bp in length
* 73522 73621: gap of unknown length
```


THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2002, 08:06:20 ; Search time 664.31 Seconds
(without alignments)
736.585 Million cell updates/sec

Title: US-09-870-113-9
Perfect score: 285
Sequence: 1 atgcagagtctacagctga.....tgaaacctgcggaagctga 285

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802.*
1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	285	100.0	1418	21	AACT74843 Human ORF398
2	280	98.2	1294	22	AAI60661 Human polynucleoti
3	280	98.2	1322	22	AAF27733 Human transport pr
4	280	98.2	1336	21	AAC90457 Human uncoupling p
5	273	95.8	401	22	AAF66156 Novel human polynu
6	272	95.4	1316	22	AAI58875 Human polynucleoti
7	225	78.9	2502	22	AA503906 Human secreted pro
8	225	78.9	27960	22	AAK69779 Human immune/haema
9	225	78.9	27960	22	AAK73320 Human immune/haema

10	174	61.1	339	22	AAK58672 Human immune/haema
11	129.6	45.5	1902	22	AAK52172 Human polynucleoti
12	128	44.9	1305	22	AAF59920 Human mitochondria
13	128	44.9	1716	21	AACT7189 Human ORF2744
14	128	44.9	2562	22	AAH90057 Human bone marrow
15	121	42.5	300	20	AAZ13026 Human gene express
16	121	42.5	710	20	AAZ15876 Human gene express
17	107.6	37.8	1913	22	AAH89944 Human bone marrow
18	105.2	36.9	483	24	ABI19871 Mouse ischaemic co
19	105	36.8	842	22	AAK92357 Human cDNA 5'-end
20	105	36.8	842	22	AAK93895 Human cDNA clone r
21	105	36.8	1814	22	AAK94871 Human full-length
22	105	36.8	2837	22	AAK80493 Human immune/haema
23	103.4	36.3	1890	22	AAK94897 Human full-length
24	103.4	36.3	3028	22	AAK91320 Human polynucleoti
25	101.8	35.7	1468	21	AACT7175 Human ORF2730
26	101	35.4	622	22	AAK91764 Human cDNA 5'-end
27	101	35.4	622	22	AAK93200 Human cDNA clone r
28	91.4	32.1	1758	20	AAV84573 Human secreted pro
29	91.4	32.1	1758	22	ABA83356 Human secreted pro
30	58	20.4	2037	24	AAK52787 cDNA sequence #574
31	42.2	14.8	2137	23	ABL08627 Drosophila melanog
32	34.4	12.1	6712	21	AACT76691 Human ORF2246
33	34	11.9	673	21	AAFI2614 Aspergillus oryzae
34	33.6	11.8	1371	21	AACT7438 Arabidopsis thalia
35	33.6	11.8	1373	21	AACT39502 Arabidopsis thalia
36	32.4	11.4	1311	18	AACT68957 Blackcurrant priB7
37	32.4	11.4	5150	18	AACT68952 Blackcurrant fruit
38	32.4	11.4	6326	23	ABL08626 Drosophila melanog
39	32	11.2	1114	24	AAI71019 Human mitochondria
40	32	11.2	1546	24	AAI71020 Human mitochondria
41	32	11.2	1560	21	AACT90452 Human uncoupling p
42	32	11.2	1662	21	AACT76992 Human ORF2547
43	31.8	11.2	51	22	AAH90263 Human clone c94399
44	31.8	11.2	51	22	AAH90264 Human clone c94399
45	31.8	11.2	878	21	AAAI5147 Probe for DNA enco

ALIGNMENTS

RESULT 1
AACT74843
ID AACT74843 standard; cDNA; 1418 BP.
AC AACT74843;
XX
XX
DT 08-FEB-2001 (first entry)
XX Human ORF398 polynucleotide sequence SEQ ID NO:795.
DE
DE
KW Human; open reading frame; ORF; detection; cytostatic; hepatotropic;
KW vulnerable; antiparasitic; antiparkinsonian; neurotropic; neuroprotective;
KW anticonvulsant; osteopathic; antithyroid; immunosuppressant; cardiac;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive; ss.
XX
XX Homo sapiens.
OS
XX
XX WO2000058473-A2.
PN
XX
XX 05-OCT-2000.
PD
XX
XX 31-MAR-2000; 2000WO-US08621.
PF

XX 31-MAR-1999; 99US-0127607.
PR 02-APR-1999; 99US-0127636.
PR 05-APR-1999; 99US-0127728.
PR 30-MAR-2000; 2000US-0540763.
XX (CURA-) CURAGEN CORP.
XX Shinkets RA, Leach M;
PI WPT: 2000-602362/57.
XX P-PSDB; AAB40634.
PT Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -
XX Claim 5; Page 853-854; 5507pp; English.
XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
XX which represent the human ORFX open reading frames 1 to 3161. The ORFX
XX sequences have activities such as: cytostatic; hepatotropic; vulnerary;
XX antiporiatic; antiparkinsonian; nootropic; neuroprotective;
XX osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
XX immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
XX antidiabetic; hypotensive; dermatological; immunosuppressive;
XX antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
XX antithyroid; and antianaemic. The sequences can be used for determining
XX the presence of or predisposition to, or preventing or treating
XX pathological conditions associated with an ORFX-associated disorder. The
XX nucleic acids can be used to express ORFX proteins in gene therapy
XX vectors. The proteins and nucleic acids may be used to treat cancers,
XX proliferative disorders, neurodegenerative disorders, osteoarthritis,
XX graft vs host disease, cardiovascular disease, diabetes mellitus, lupus
XX hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
XX erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
XX bacterial or fungal infection, malaria, autoimmune disorders, asthma,
XX allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
XX nocturnal haemoglobinuria, antiinflammatory disease; to enhance
XX coagulation; to inhibit thrombosis; and as a contraceptive.
XX Sequence 1418 BP; 353 A; 400 C; 368 G; 297 T; 0 other;

Query Match 100.0%; Score 285; DB 21; Length 1418;
Best Local Similarity 100.0%; Pred. No. 9.3e-87;
Matches 285; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 atgcagagttctacagcctgacccagctgccgctatcgcaatgtgttgaggccctctgg 60
Db 178 atgcagagttctacagcctgacccagctgccgctatcgcaatgtgttgaggccctctgg 237
Qy 61 aggtattataagaacgaggcctatgagcccatgagggcgtgacgtcacagcaaca 120
Db 238 aggtattataagaacgaggcctatgagcccatgagggcgtgacgtcacagcaaca 297
Qy 121 ggcgcagggcctccacgcctttatttgcctgctacgaaagttaaaaaagacattg 180
Db 298 ggcgcagggcctccacgcctttatttgcctgctacgaaagttaaaaaagacattg 357
Qy 181 agtgatgaatccacctgggggcaatagccatttgcctgctacgaaagttaaaaaagacattg 240
Db 358 agtgatgaatccacctgggggcaatagccatttgcctgctacgaaagttaaaaaagacattg 417
Qy 241 gcaacattacttcatgatgcagccatgaacccctgcgggaagcctga 285
Db 418 gcaacattacttcatgatgcagccatgaacccctgcgggaagcctga 462

RESULT 2
AAI60661/c
ID AAI60661 standard; cDNA; 1294 BP.
XX

AC AAI60661;
XX 22-OCT-2001 (first entry)
XX Human polynucleotide SEQ ID NO 4650.
XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX Homo sapiens.
XX WO200153312-A1.
XX 26-JUL-2001.
XX 26-DEC-2000; 2000WO-US34263.
XX 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-052317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 14-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX WPI: 2001-442253/47.
DR P-PSDB; AAM41505.
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX Claim 1; SEQ ID NO 4650; 10078pp; English.
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX Sequence 1294 BP; 267 A; 335 C; 368 G; 324 T; 0 other;

Query Match 98.2%; Score 280; DB 22; Length 1294;
Best Local Similarity 100.0%; Pred. No. 4.5e-85;
Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 atgcagagttctacagcctgacccagctgccgctatcgcaatgtgttgaggccctctgg 60
Db 1134 ATGCAGAGTCTACAGCCTGACCCAGCTGCCGCTATCGCAATGTGTGAGGCCCTCTGG 1075
Qy 61 aggtattataagaacgaggcctatgagcccatgagggcgtgacgtcacagcaaca 120

Db	1074	AGGATTATAAGAACGGAGGGCTATGGAGGCCCATGAGGGGCTGAACGTCACAGAAACA	1015
Qy	121	ggcgagggcctgccacgcctttatttgcctgctacgaaaaagttaaaaaagacattg	180
Db	1014	GGCGAGGGCTGCCACGCCCTTTATTTTGCTGCTACTGAAAAAGTTAAAAAAGACATTG	955
Qy	181	agtgatgtaatccaccctgggggccaatagccatattgccaatggtgcgcccgggtgtg	240
Db	954	AGTGATGTAATCACCTCGGGGCAATAGCCATATTGCCAATGTTGCGCGCGGTGTGTG	895
Qy	241	gcaacattacttcattgatgcacgcacatgaacctgcggaag	280
Db	894	GCAACATTACTTCATGATGCAGCCATGAACCCCTGCGGAAG	855
RESULT	3		
ID	AAF27733		
XX	AAF27733	standard; cDNA; 1322 BP.	
AC	AAF27733;		
XX			
DT	28-MAR-2001	(first entry)	
XX			
DE	Human transport protein TPPT-33	coding sequence.	
XX			
KW	Human; transport protein; TPPT; transport disorder; metabolic disorder;		
KW	neurological disorder; cardiovascular disorder; reproductive disorder;		
KW	immune disorder; cancer; ss.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200078953-A2.		
XX			
PD	28-DEC-2000.		
XX			
PF	16-JUN-2000; 2000MO-US16668.		
XX			
PR	17-JUN-1999; 99US-0139923.		
PR	10-AUG-1999; 99US-0148177.		
PR	18-AUG-1999; 99US-0149357.		
PR	28-OCT-1999; 99US-0162287.		
XX			
PA	(INCY-) INCYTE GENOMICS INC.		
XX			
PI	Lal P, Yang J, Yue H, Hillman JL, Tang YT, Bandman O, Burford N;		
PI	Baughn MR, Azimzai Y, Lu DAM, Au-Young J, Patterson C;		
XX			
DR	WPI; 2001-041424/05.		
DR	P-PSDB; AAB60113.		
XX			
PT	Isolated polypeptide with a human transport protein sequence is useful		
PT	for the diagnosis, prevention and treatment of disorders associated		
PT	with the immune, reproductive and cardiovascular systems -		
XX			
PS	Claim 5; Page 160; 165pp; English.		
XX			
CC	The present invention provides the protein and coding sequences for 43		
CC	novel human transport proteins (designated TPPTs). These can be used in		
CC	the diagnosis and treatment of transport, metabolic, neurological,		
CC	reproductive, cardiovascular and immune disorders, and cell proliferative		
CC	disorders such as cancer.		
XX			
SQ	Sequence 1322 BP; 328 A; 376 C; 351 G; 267 T; 0 other;		
Query Match	98.2%;	Score 280;	DB 22; Length 1322;
Best local Similarity	100.0%;	Pred. No. 4.6e-85;	
Matches 280;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
Qy	1	atgcagaggtctacagcctgacccagctgcgcgctatgcacatgtgttgaggccctctgg	60
Db	186	atgcagaggtctacagcctgacccagctgcgcgctatgcacatgtgttgaggccctctgg	245

```
Query Match      98.2%; Score 280; DB 21; Length 1336;
Best Local Similarity 100.0%; Pred. No. 4.6e-85;
Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atcgagctctacagcctgacccagctcccgctatcgcaatgtgttgaggccctctgg 60
Db 197 atcgagctctacagcctgacccagctcccgctatcgcaatgtgttgaggccctctgg 256

QY 61 aggtattataagaacgaggcctatgaggccatgaggggctgaacgtcacagcaaca 120
Db 257 aggtattataagaacgaggcctatgaggccatgaggggctgaacgtcacagcaaca 316

QY 121 ggcgcaggcctgcccacgcccctttatttgcctgctacgaaaagttaaaaaagacattg 180
Db 317 ggcgcaggcctgcccacgcccctttatttgcctgctacgaaaagttaaaaaagacattg 376

QY 181 agtgaatgaatccaccctgggggcaatagccatattgccaatgtgtgcggcggtgtgtg 240
Db 377 agtgaatgaatccaccctgggggcaatagccatattgccaatgtgtgcggcggtgtgtg 436

QY 241 gcaacattactctatgatgcagccatgaacccctgcggaag 280
Db 437 gcaacattactctatgatgcagccatgaacccctgcggaag 476

RESULT 5
AAF66156
ID AAF66156 standard; cDNA; 401 BP.
XX
AC AAF66156;
XX
DT 09-APR-2001 (first entry)
XX
DE Novel human polynucleotide, SEQ ID NO: 1912.
XX
KW Human; cytostatic; gene therapy; colon cancer; prostate cancer;
KW breast cancer; lung cancer; cancer detection; ss.
XX
OS Homo sapiens.
XX
PN WO200102568-A2.
XX
PD 11-JAN-2001.
XX
PF 30-JUN-2000; 2000WO-US18374.
XX
PR 02-JUL-1999; 99US-0142310.
XX
PR 02-JUL-1999; 99US-0142311.
XX
PA (CHIR ) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
PI Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A;
PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G, Drmanac R;
PI Crkenjakov R, Drmanac S, Dickson M, Labat I, Leshkowitz D;
PI Kita D, Garcia V, Jones LW, Strache-Crain B;
XX
DR WPI; 2001-091805/10.
XX
PT Library of polynucleotides for diagnosing a cancerous state of a
PT mammalian cell and detecting cancer, particularly of the colon or
PT prostate, comprises 3351 human polynucleotide sequences -
XX
PS Claim 9; Page 818; 1046pp; English.
XX
CC The present sequence is one of 3351 sequences in a library of human
CC polynucleotides. The library is used to detect differentially expressed
CC genes correlated with a cancerous state of a mammalian cell and can
CC detect colon, prostate, breast and lung cancer. The library can be used
CC to produce probes for detection of mRNA and to produce additional copies
CC of the polynucleotides. The probes can be used for chromosome mapping of
CC the polynucleotide and for detection of transcription levels. Ribozymes
```

```
CC or antisense oligonucleotides can be generated. The polynucleotides and
CC their gene products are used as genetic or biochemical markers (e.g. in
CC blood or tissues) that will detect the earliest changes along the
CC carcinogenesis pathway and/or monitor the efficacy of therapies and
CC preventive interventions. The polynucleotides, polypeptides and
CC antibodies against them can be used in pharmaceutical compositions to
CC treat the cancers and proliferative disorders such as neoplasia,
CC dysplasia and hyperplasia.
XX
SQ Sequence 401 BP; 100 A; 104 C; 120 G; 76 T; 1 other;

Query Match      95.8%; Score 273; DB 22; Length 401;
Best Local Similarity 100.0%; Pred. No. 6.5e-83;
Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 gtctacagcctgacccagctcccgctatcgcaatgtgttgaggccctctggaggatta 67
Db 9 gtctacagcctgacccagctcccgctatcgcaatgtgttgaggccctctggaggatta 68

QY 68 taagaacgaggggcctatgaggcccatgaggggctgaacgtcacagcaacagcgcgag 127
Db 69 taagaacgaggggcctatgaggcccatgaggggctgaacgtcacagcaacagcgcgag 128

QY 128 ggcctgccacgcccctttatttgcctgctacgaaaagttaaaaaagacattgagtgtg 187
Db 129 ggcctgccacgcccctttatttgcctgctacgaaaagttaaaaaagacattgagtgtg 188

QY 188 taatccaccctgggggcaatagccatattgccaatgtgtgcggcggtgtgtgcaacat 247
Db 189 taatccaccctgggggcaatagccatattgccaatgtgtgcggcggtgtgtgcaacat 248

QY 248 tacttcatgatgcagccatgaacccctgcggaag 280
Db 249 tacttcatgatgcagccatgaacccctgcggaag 281

RESULT 6
AAI58875
ID AAI58875 standard; cDNA; 1316 BP.
XX
AC AAI58875;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 1078.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
XX
PR 25-APR-2000; 2000US-0552317.
XX
PR 09-JUL-2000; 2000US-0598042.
XX
PR 19-JUL-2000; 2000US-0620312.
XX
PR 03-AUG-2000; 2000US-0653450.
XX
PR 14-SEP-2000; 2000US-0662191.
XX
PR 19-OCT-2000; 2000US-0693036.
XX
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
```


XX	06-NOV-2001	(first entry)
DT	Human immune/haematopoietic antigen genomic sequence	SEQ ID NO:24591.
XX	Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;	
DE	cytostatic; gene therapy; vaccine; metastasis; ds.	
XX	Homo sapiens.	
KW	WO200157182-A2.	
XX	09-AUG-2001.	
OS	17-JAN-2001; 2001WO-US01354.	
PN	31-JAN-2000; 2000US-0179065.	
XX	04-FEB-2000; 2000US-0180628.	
PR	24-FEB-2000; 2000US-0184664.	
XX	02-MAR-2000; 2000US-0186350.	
PR	16-MAR-2000; 2000US-0189874.	
XX	17-MAR-2000; 2000US-0190076.	
PR	18-APR-2000; 2000US-0198123.	
XX	19-MAY-2000; 2000US-0205515.	
PR	07-JUN-2000; 2000US-0209467.	
XX	28-JUN-2000; 2000US-0214886.	
PR	30-JUN-2000; 2000US-0215135.	
XX	07-JUL-2000; 2000US-0216647.	
PR	07-JUL-2000; 2000US-0216880.	
XX	11-JUL-2000; 2000US-0217487.	
PR	11-JUL-2000; 2000US-0217496.	
XX	14-JUL-2000; 2000US-0218290.	
PR	26-JUL-2000; 2000US-0220963.	
XX	26-JUL-2000; 2000US-0220964.	
PR	14-AUG-2000; 2000US-0224518.	
XX	14-AUG-2000; 2000US-0224519.	
PR	14-AUG-2000; 2000US-0225213.	
XX	14-AUG-2000; 2000US-0225214.	
PR	14-AUG-2000; 2000US-0225266.	
XX	14-AUG-2000; 2000US-0225267.	
PR	14-AUG-2000; 2000US-0225268.	
XX	14-AUG-2000; 2000US-0225270.	
PR	14-AUG-2000; 2000US-0225447.	
XX	14-AUG-2000; 2000US-0225757.	
PR	14-AUG-2000; 2000US-0225758.	
XX	14-AUG-2000; 2000US-0225759.	
PR	18-AUG-2000; 2000US-0226279.	
XX	22-AUG-2000; 2000US-0226681.	
PR	22-AUG-2000; 2000US-0226868.	
XX	23-AUG-2000; 2000US-0227182.	
PR	23-AUG-2000; 2000US-0227009.	
XX	30-AUG-2000; 2000US-0228924.	
PR	01-SEP-2000; 2000US-0229287.	
XX	01-SEP-2000; 2000US-0229343.	
PR	01-SEP-2000; 2000US-0229344.	
XX	01-SEP-2000; 2000US-0229345.	
PR	05-SEP-2000; 2000US-0229509.	
XX	05-SEP-2000; 2000US-0229513.	
PR	06-SEP-2000; 2000US-0230437.	
XX	06-SEP-2000; 2000US-0230438.	
PR	08-SEP-2000; 2000US-0231242.	
XX	08-SEP-2000; 2000US-0231243.	
PR	08-SEP-2000; 2000US-0231244.	
XX	08-SEP-2000; 2000US-0231413.	
PR	08-SEP-2000; 2000US-0231414.	
XX	08-SEP-2000; 2000US-0232080.	
PR	08-SEP-2000; 2000US-0232081.	
XX	12-SEP-2000; 2000US-0231968.	
PR	14-SEP-2000; 2000US-0232397.	
XX	14-SEP-2000; 2000US-0232398.	
PR	14-SEP-2000; 2000US-0232399.	
XX	14-SEP-2000; 2000US-0232401.	
PR	14-SEP-2000; 2000US-0232401.	

PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis.
XX
XX
PS Disclosure; SEQ ID NO 24591; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention.
XX
XX Sequence 27960 BP; 7998 A; 5835 C; 6583 G; 7544 T; 0 other;
SQ

Query Match 78.9%; Score 225; DB 22; Length 27960;
Best Local Similarity 100.0%; Pred. No. 9.8e-66;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgcagagtctacagcctgaccagctgccgctatcgcaatgtgtggagccctctgg 60
|||||
Db 18295 atgcagagtctacagcctgaccagctgccgctatcgcaatgtgtggagccctctgg 18354
|||||

QY 61 aggattataagaacgagggcctatggagcccatgaggggctgaacgtcacagcaaca 120
|||||
Db 18355 aggattataagaacgagggcctatggagcccatgaggggctgaacgtcacagcaaca 18414
|||||

QY 121 ggcgcagggcctgccacgcccctttatttgcgtgtacgaaagttaaaaagacattg 180
|||||
Db 18415 ggcgcagggcctgccacgcccctttatttgcgtgtacgaaagttaaaaagacattg 18474
|||||

QY 181 agtgcgttaataccacctgggggcaatagccattatggccaatgg 225
|||||
Db 18475 agtgcgttaataccacctgggggcaatagccattatggccaatgg 18519
|||||

RESULT 9
AAK73320
ID AAK73320 standard; DNA; 27960 BP.
XX
AC AAK73320;
XX
XX 06-NOV-2001 (first entry)
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:28132.
DE Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX cytostatic; gene therapy; vaccine; metastasis; ds.
KW

XX Homo sapiens.
OS WO200157182-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 17-JAN-2001; 2001WO-US01354.
PF
XX
XX 31-JAN-2000; 2000US-0179065.
XX 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0228868.
PR 23-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.

PF	17-JAN-2001;	2001WO-US01354.	
XX	31-JAN-2000;	2000US-0179065.	
PR	04-FEB-2000;	2000US-0180628.	
PR	24-FEB-2000;	2000US-0184664.	
PR	02-MAR-2000;	2000US-0186350.	
PR	16-MAR-2000;	2000US-0189874.	
PR	17-MAR-2000;	2000US-0190076.	
PR	18-APR-2000;	2000US-0198123.	
PR	19-MAY-2000;	2000US-0205515.	
PR	07-JUN-2000;	2000US-0209467.	
PR	28-JUN-2000;	2000US-0214886.	
PR	30-JUN-2000;	2000US-0215135.	
PR	07-JUL-2000;	2000US-0216647.	
PR	07-JUL-2000;	2000US-0216680.	
PR	11-JUL-2000;	2000US-0217487.	
PR	11-JUL-2000;	2000US-0217496.	
PR	14-JUL-2000;	2000US-0218290.	
PR	26-JUL-2000;	2000US-0220963.	
PR	26-JUL-2000;	2000US-0220964.	
PR	14-AUG-2000;	2000US-0224518.	
PR	14-AUG-2000;	2000US-0224519.	
PR	14-AUG-2000;	2000US-0225213.	
PR	14-AUG-2000;	2000US-0225214.	
PR	14-AUG-2000;	2000US-0225266.	
PR	14-AUG-2000;	2000US-0225267.	
PR	14-AUG-2000;	2000US-0225268.	
PR	14-AUG-2000;	2000US-0225270.	
PR	14-AUG-2000;	2000US-0225447.	
PR	14-AUG-2000;	2000US-0225757.	
PR	14-AUG-2000;	2000US-0225758.	
PR	14-AUG-2000;	2000US-0225759.	
PR	18-AUG-2000;	2000US-0226279.	
PR	22-AUG-2000;	2000US-0226681.	
PR	22-AUG-2000;	2000US-0226688.	
PR	22-AUG-2000;	2000US-0227182.	
PR	23-AUG-2000;	2000US-0227009.	
PR	30-AUG-2000;	2000US-0228924.	
PR	01-SEP-2000;	2000US-0229287.	
PR	01-SEP-2000;	2000US-0229343.	
PR	01-SEP-2000;	2000US-0229344.	
PR	01-SEP-2000;	2000US-0229345.	
PR	05-SEP-2000;	2000US-0229509.	
PR	05-SEP-2000;	2000US-0229513.	
PR	06-SEP-2000;	2000US-0230437.	
PR	06-SEP-2000;	2000US-0230438.	
PR	08-SEP-2000;	2000US-0231242.	
PR	08-SEP-2000;	2000US-0231243.	
PR	08-SEP-2000;	2000US-0231244.	
PR	08-SEP-2000;	2000US-0231413.	
PR	08-SEP-2000;	2000US-0231414.	
PR	08-SEP-2000;	2000US-0232080.	
PR	08-SEP-2000;	2000US-0232081.	
PR	12-SEP-2000;	2000US-0231968.	
PR	14-SEP-2000;	2000US-0232397.	
PR	14-SEP-2000;	2000US-0232398.	
PR	14-SEP-2000;	2000US-0232399.	
PR	14-SEP-2000;	2000US-0232400.	
PR	14-SEP-2000;	2000US-0232401.	
PR	14-SEP-2000;	2000US-0233063.	
PR	14-SEP-2000;	2000US-0233064.	
PR	14-SEP-2000;	2000US-0233065.	
PR	21-SEP-2000;	2000US-0234223.	
PR	21-SEP-2000;	2000US-0234274.	
PR	25-SEP-2000;	2000US-0234997.	
PR	25-SEP-2000;	2000US-0234998.	
PR	26-SEP-2000;	2000US-0235484.	
PR	27-SEP-2000;	2000US-0235834.	
PR	27-SEP-2000;	2000US-0235836.	
PR	29-SEP-2000;	2000US-0236327.	
PR	29-SEP-2000;	2000US-0236367.	
PR	29-SEP-2000;	2000US-0236368.	
PR	29-SEP-2000;	2000US-0236369.	
PR	29-SEP-2000;	2000US-0236370.	
PR	02-OCT-2000;	2000US-0236802.	
PR	02-OCT-2000;	2000US-0237037.	
PR	02-OCT-2000;	2000US-0237038.	
PR	02-OCT-2000;	2000US-0237039.	
PR	02-OCT-2000;	2000US-0237040.	
PR	13-OCT-2000;	2000US-0239935.	
PR	13-OCT-2000;	2000US-0239937.	
PR	20-OCT-2000;	2000US-0240960.	
PR	20-OCT-2000;	2000US-0241221.	
PR	20-OCT-2000;	2000US-0241785.	
PR	20-OCT-2000;	2000US-0241786.	
PR	20-OCT-2000;	2000US-0241787.	
PR	20-OCT-2000;	2000US-0241808.	
PR	20-OCT-2000;	2000US-0241809.	
PR	20-OCT-2000;	2000US-0241826.	
PR	01-NOV-2000;	2000US-0244617.	
PR	08-NOV-2000;	2000US-0244674.	
PR	08-NOV-2000;	2000US-0246475.	
PR	08-NOV-2000;	2000US-0246476.	
PR	08-NOV-2000;	2000US-0246477.	
PR	08-NOV-2000;	2000US-0246478.	
PR	08-NOV-2000;	2000US-0246523.	
PR	08-NOV-2000;	2000US-0246524.	
PR	08-NOV-2000;	2000US-0246525.	
PR	08-NOV-2000;	2000US-0246526.	
PR	08-NOV-2000;	2000US-0246527.	
PR	08-NOV-2000;	2000US-0246528.	
PR	08-NOV-2000;	2000US-0246532.	
PR	08-NOV-2000;	2000US-0246609.	
PR	08-NOV-2000;	2000US-0246610.	
PR	08-NOV-2000;	2000US-0246611.	
PR	08-NOV-2000;	2000US-0246613.	
PR	17-NOV-2000;	2000US-0249207.	
PR	17-NOV-2000;	2000US-0249208.	
PR	17-NOV-2000;	2000US-0249209.	
PR	17-NOV-2000;	2000US-0249210.	
PR	17-NOV-2000;	2000US-0249211.	
PR	17-NOV-2000;	2000US-0249212.	
PR	17-NOV-2000;	2000US-0249213.	
PR	17-NOV-2000;	2000US-0249214.	
PR	17-NOV-2000;	2000US-0249215.	
PR	17-NOV-2000;	2000US-0249216.	
PR	17-NOV-2000;	2000US-0249217.	
PR	17-NOV-2000;	2000US-0249218.	
PR	17-NOV-2000;	2000US-0249244.	
PR	17-NOV-2000;	2000US-0249245.	
PR	17-NOV-2000;	2000US-0249264.	
PR	17-NOV-2000;	2000US-0249265.	
PR	17-NOV-2000;	2000US-0249297.	
PR	17-NOV-2000;	2000US-0249299.	
PR	17-NOV-2000;	2000US-0249300.	
PR	01-DEC-2000;	2000US-0250160.	
PR	01-DEC-2000;	2000US-0250391.	
PR	05-DEC-2000;	2000US-0251030.	
PR	05-DEC-2000;	2000US-0251988.	
PR	05-DEC-2000;	2000US-0256719.	
PR	06-DEC-2000;	2000US-0251479.	
PR	08-DEC-2000;	2000US-0251856.	
PR	08-DEC-2000;	2000US-0251868.	
PR	08-DEC-2000;	2000US-0251869.	
PR	08-DEC-2000;	2000US-0251989.	
PR	08-DEC-2000;	2000US-0251990.	
PR	11-DEC-2000;	2000US-0254097.	
PR	05-JAN-2001;	2001US-0259678.	
XX	(HUMA-)	HUMAN GENOME SCI INC.	
PA	Rosen CA,	Barash SC,	Ruben SM;
PI	WPI;	2001-483426/52.	
DR	P-PSDB;	AAM85891.	
DR			
XX			

PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
XX metastasis -
PS Claim 1: SEQ ID NO 3732; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 339 BP; 86 A; 84 C; 101 G; 64 T; 4 other;

Query Match 61.18; Score 174; DB 22; Length 339;
Best Local Similarity 94.98; Pred. No. 3.3e-49;
Matches 188; Conservative 1; Mismatches 8; Indels 1; Gaps 1;
Qy 1 atgcagagctctacagcctgacccagctgcccgcctatgcgaatgtgtgagggccctctgg 60
Db 143 atgcagwtctacagctgacccagctgcccgcctatgcgaatgtgtgagggccctctgg 202
Qy 61 aggattataagaacgaggggcccctatggagggccatgaggggctgaaactcacagcaaca 120
Db 203 aggattataagaacgaggggcccctatggagg-ccatgagggggctgaaactcacagcaaca 261
Qy 121 ggcgcagggcctgccacgcccctttatttgcctgctacgaaagttaaaaaagacattg 180
Db 262 ggcgcagggcctnccacgcccctttatttgcctgctacgaaagttaaaaaagacattg 321
Qy 181 agtgatgtaataccacct 198
Db 322 agtgatgtaataccacct 339

RESULT 11
AAK52172
ID AAK52172 standard; cDNA; 1902 BP.
AC AAK52172;
XX
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 717.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
PN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US04098.
XX
XX 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.

PR 20-JUN-2000; 2000US-0598075.
PR 19-JUL-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0663561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
DR WPI: 2001-476283/51.
DR P-PSDB; AAM79039.
XX
PT Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
XX
PS Claim 1: Page 2450-2452; 6221pp; English.
XX
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAM80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.
XX
SQ Sequence 1902 BP; 434 A; 532 C; 547 G; 389 T; 0 other;

Query Match 45.5%; Score 129.6; DB 22; Length 1902;
Best Local Similarity 66.4%; Pred. No. 9.6e-34;
Matches 186; Conservative 0; Mismatches 94; Indels 0; Gaps 0;
Qy 1 atgcagagctctacagcctgacccagctgcccgcctatgcgaatgtgtgagggccctctgg 60
Db 1278 atgcagagcttgcagctccagctcccaagcccaagctacacaaagtgtctacggccctcaag 1337
Qy 61 aggattataagaacgaggggcccctatggagggccatgaggggctgaaactcacagcaaca 120
Db 1338 aaatcatcgagcgaagggtcttgcggcccttgcggcctgcacacgtcatgatcatg 1397
Qy 121 ggcgcagggcctgccacgcccctttatttgcctgctacgaaagttaaaaaagacattg 180
Db 1398 ggtgcagggcgcagcccatgcccattgttttgcctgctatgaaacatgaaaggacttta 1457
Qy 181 agtgatgtaataccacctgggggccaatagccatattgccaatgtgtgcggccgggtgtg 240
Db 1458 aatgacgttttccaccaccaagaaacagccacctagccaagggatagctggggagtatg 1517
Qy 241 gcaacattactcatgatgagcgcctatgaaccttcgggaag 280
Db 1518 gccacctgctccacgctgcggtaatgaatccacgacgaag 1557

RESULT 12
AAF59920
ID AAF59920 standard; cDNA; 1305 BP.
XX
AC AAF59920;
XX
XX
DT 04-MAY-2001 (first entry)
XX
DE Human mitochondrial solute carrier protein hMSC-o cDNA.

Db 316 aaatcatgcggaccgagcctctctgaggcccttgagcgctcaacgtcatgatcatg 375
Qy 121 ggcagagcctgccacgccttatttgcctgtactacgaaagttaaaaaagacattg 180
Db 376 ggtcagggccagccatgcctatgtatttgcctgtatgaaacatgaaaggacttta 435
Qy 181 agtcatgaatccaccctgggggcaatagccatatattgccaatggtgcggcggtgtgtg 240
Db 436 aatgacgttttccaccacgaagaaacagccacctagccaacggatagctgggagtatg 495
Qy 241 gcaacattactcatgatgcagcccatgaacccctgcggaag 280
Db 496 gccacctgtcccagatgcggtgaatgaatccagcagaag 535

RESULT 14
AAH90057
ID AAH90057 standard; cDNA; 2562 BP.
XX AC
XX AAH90057;
XX
DT 01-OCT-2001 (first entry)
XX
XX Human bone marrow cDNA, SEQ ID NO: 301.
XX
XX Human; bone marrow; antiinflammatory; cytostatic; neuroprotective;
KW antiviral; antibacterial; antifungal; anti-HIV; haemostatic;
KW immunosuppressive; gene therapy; cytokine cell proliferation;
KW cell differentiation modulator; immune disorder; infection; cancer;
KW human immunodeficiency virus; HIV; autoimmune disorder; haemophilia; ss.
XX
XX Homo sapiens.
XX
XX WO200153453-A2.
XX
PD 26-JUL-2001.
XX
XX
XX 23-DEC-2000; 2000WO-US34960.
XX
XX 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 30-NOV-2000; 2000US-0250583.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;
PI Ren F, Wang J, Werhman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Drmanac RT;
XX
XX WPI: 2001-488707/53.
DR P-PSDB; AAM00938.
XX
XX Novel bone-marrow-expressed polynucleotides and polypeptides, useful
PT for treating e.g. cancer and immune deficiency disorders -
XX
XX Claim 1; Page 412-413; 648pp; English.
XX
XX The present sequence is one of 251 novel human polynucleotides
CC expressed in the bone marrow. The polynucleotide and the
CC polypeptide encoded by it are useful in the treatment of various
CC immune deficiencies and disorders. The deficiencies and disorders may
CC be genetic, may be caused by a viral (e.g. HIV), bacterial or fungal
CC infection, or may result from an autoimmune disorder, a coagulation
CC disorder (e.g. haemophilia), inhibition of tumour cell proliferation,
CC suppression of an inflammatory response or treatment of a nervous
CC system disorder such as Alzheimer's disease. Detection of the presence
CC or increased expression of the polynucleotide or the protein it
CC encodes is useful for the diagnosis and/or prognosis of one

CC or more types of cancer. The polynucleotide and polypeptide can be
CC used as nutritional sources or supplements and in the screening of
CC chemical compounds as potential drugs.
XX
SQ Sequence 2562 BP; 623 A; 660 C; 627 G; 652 T; 0 other;

Query Match 44.9%; Score 128; DB 22; Length 2562;
Best Local Similarity 66.1%; Pred. No. 3.8e-33;
Matches 185; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

Qy 1 atgcagagtctacagcctgaccagctgcgcgtctatcgcaatgtgttgaggccctctgg 60
Db 2021 atgcagagtttgagtcacagatcccaagccagctacacagatctctacggagccctcaag 2080
Qy 61 aggattataagacggagggcctatgagcccatgaggggctgaacgtcacagcaaca 120
Db 2081 aaatcatgcggaccgagcctctctgagggccttgcagcctgcacgtcatgatcatg 2140
Qy 121 ggcagggcctgccacgccttatttgcctgtactacgaaagttaaaaaagacattg 180
Db 2141 ggtcagggcggcccatgcccattgtatttgcctgtatgaaacatgaaaggacttta 2200
Qy 181 agtcatgaatccaccctgggggcaatagccatatattgccaatggtgcggcggtgtgtg 240
Db 2201 aatgacgttttccaccacgaagaaacagccacctagccaacggatagctgggagtatg 2260
Qy 241 gcaacattactcatgatgcagcccatgaacccctgcggaag 280
Db 2261 gccacctgtcccagatgcggtgaatgaatccagcagaag 2300

RESULT 15
AAZ13026
ID AAZ13026 standard; cDNA; 300 BP.
XX AC
XX AAZ13026;
XX
DT 12-OCT-1999 (first entry)
XX
XX Human gene expression product cDNA sequence SEQ ID NO:495.
XX
XX Human; gene; gene expression product; diagnosis; therapy; probe;
KW detection; mapping; tissue typing; profiling; forensic; cancer;
KW genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.
XX
XX Homo sapiens.
XX
XX WO9938972-A2.
XX
PD 05-AUG-1999.
XX
XX 28-JAN-1999; 99WO-US01619.
XX
XX 03-APR-1998; 98US-0080666.
PR 28-JAN-1998; 98US-0072910.
PR 24-FEB-1998; 98US-0075954.
PR 31-MAR-1998; 98US-0080114.
PR 03-APR-1998; 98US-0080515.
XX
XX (CHIR) CHIRON CORP.
XX (HYSE-) HYSEQ INC.
XX
XX Crkvenjakov R, Dickson M, Drmanac R, Drmanac S;
PI Escobedo J, Garcia PD, Garcia V, Giese K, Innis MA;
PI Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;
PI Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;
PI Stache-Crain B, Sudduth-Klinger J, Williams LT;
XX
XX WPI: 1999-494092/41.
DR
XX Novel human genes and their expression products which are
PT differentially expressed in different cell types

Search completed: September 28, 2002, 08:06:28
Job time: 11583 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2002, 07:55:54 ; Search time 182.06 Seconds
(without alignments)
384.519 Million cell updates/sec

Title: US-09-870-113-9
Perfect score: 285
Sequence: 1 atcgagtgctacagctga.....tgaaacctggcgaagctga 285

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	32.4	11.4	1311	4	US-09-068-140A-9
2	32.4	11.4	1510	4	US-09-068-140A-14
3	32	11.2	1643	2	US-08-933-750C-68
4	32	11.2	1643	3	US-09-234-613-68
5	29.6	10.4	4091	4	US-08-630-915A-33
6	29	10.2	1747	1	US-08-202-389-1
7	29	10.2	8600	4	US-09-457-708-1
8	28.4	10.0	4368	1	US-08-026-138E-17
9	28.4	10.0	4446	1	US-08-026-138E-6
10	28.2	9.9	7011	4	US-09-268-163-9
11	28	9.8	2379	3	US-08-537-361E-5
12	27.6	9.7	3942	4	US-09-162-484-19
13	27.2	9.5	3771	1	US-08-185-432-3
14	27.2	9.5	3771	1	US-08-185-432-23
15	27.2	9.5	3810	2	US-08-475-844-8
16	27.2	9.5	3810	5	PCT-US95-08429-8
17	27.2	9.5	5063	1	US-08-185-432-1
18	27	9.5	1710	3	US-09-106-182-1
19	27	9.5	5087	1	US-08-792-019B-3
20	27	9.5	5087	3	US-08-988-819-3
21	27	9.5	5087	3	US-09-016-534-3
22	27	9.5	7653	4	US-08-471-112A-1
23	27	9.5	7824	5	PCT-US95-06722-11
24	26.8	9.4	1192	4	US-09-142-565-1
25	26.8	9.4	1255	1	US-08-518-878B-38
26	26.8	9.4	1255	1	US-08-294-522B-38
27	26.8	9.4	1255	2	US-08-470-868A-38

28	26.8	9.4	1596	2	US-08-807-861A-38	Sequence 38, Appl
29	26.8	9.4	1596	3	US-09-210-681-38	Sequence 38, Appl
30	26.8	9.4	1596	3	US-08-946-719A-38	Sequence 38, Appl
c 31	26.8	9.4	3672	4	US-09-056-105-17	Sequence 17, Appl
32	26.8	9.4	246240	2	US-08-724-394A-20	Sequence 20, Appl
33	26.8	9.4	246240	2	US-08-724-394A-21	Sequence 21, Appl
34	26.8	9.4	246240	2	US-08-724-394A-22	Sequence 22, Appl
c 35	26.6	9.3	2130	1	US-07-952-800-1	Sequence 1, Appl
36	26.6	9.3	4692	2	US-08-916-917-1	Sequence 1, Appl
37	26.6	9.3	4692	2	US-08-972-631-1	Sequence 1, Appl
38	26.6	9.3	4692	2	US-08-972-629-1	Sequence 1, Appl
39	26.6	9.3	4692	2	US-08-972-630-1	Sequence 1, Appl
40	26.6	9.3	4692	2	US-08-672-211-1	Sequence 1, Appl
41	26.6	9.3	4692	3	US-09-225-170-1	Sequence 1, Appl
42	26.4	9.3	1777	2	US-08-937-466-5	Sequence 5, Appl
43	26.4	9.3	1777	2	US-09-172-528-5	Sequence 5, Appl
44	26.4	9.3	1777	3	US-09-318-199-5	Sequence 5, Appl
45	26.4	9.3	1777	4	US-09-503-579-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-09-068-140A-9
; Sequence 9, Application US/09068140A
; Patent No. 6281409
; GENERAL INFORMATION:
; APPLICANT: Mary Rose Woodhead, Mark Andrew Taylor
; APPLICANT: and Rex Michael Brennan
; TITLE OF INVENTION: Blackcurrant Promoters and Genes
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/068.140A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP96/04807
; FILING DATE: No. 6281409ember 4, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Dinner, Dara L.
; REGISTRATION NUMBER: 33,680
; REFERENCE/DOCKET NUMBER: C70237
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5017
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1311 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: cdna
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Ribes nigrum
; STRAIN: Ben Alder
US-09-068-140A-9

Db	482	CTACCCGCTTCAC	TGGCACCAT	TGATGCC	TTCGTGAAGAT	CTGTGAGGCACGAGGGCACCA	541
Qy	86	ggagggccatga	gggggct	aaagtc	acagcaac	gagcgagggccctgcccagcgcccttt	145
Db	542	GGACCCCTCTG	TGAGGGGCT	CCCCGCC	ACCCTGTGTGAT	GACTGTGCCAGTACCGCCGATCT	601
Qy	146	atttgcctgctac	gaaaaagt	ataaaa	gacatt	gagtgatgtaatcc	193
Db	602	ACTTCATGCTC	TATGACCA	CTGAAGGCC	TTCTGTGTGCTGCAGGCC		649

```

RESULT      4
US-09-234-613-68
; Sequence 68, Application US/09234613
; Patent No. 6132973
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/234,613
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,750
; FILING DATE: September 23, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0356 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1643 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: SYN00AT01
; CLONE: 724157
US-09-234-613-68

```

	Query Match	11.28;	Score 32;	DB 3;	Length 1643;	
	Best Local Similarity	49.4%;	Pred. No.	0.16;		
	Matches	83;	Conservative	0;	Mismatches	85; Indels 0; Gaps 0;
Qy	26	ctgcgccgctatcgcaattgttggaggccctctcggaggaattataagaacgagagggccttat	85			
Db	482	ctaccgcctttcacrtgggacacattgatgcctttcgttgagatcgtgaggcacggaggcaccca	541			

```

Qy      86  ggagggccatgaggggctgaacgtcacagcaacagggcagggctgccccagcccttt 145
Db      542  GGACCCTCTGGAGCGGCTCCCCCGCCACCTGGTGTGATGACTGTGCGAGCTACCGCATCT 601
Qy      146  atttgctgctcagaaaagttaaaaaagacattgagtgtgtaattcc 193
Db      602  ACTTCATGCTATGACCACTGAAGCCCTTCTCTGTGTGGTCGAGCCC 649

RESULT      5
US-08-630-915A-33
; Sequence 33, Application US/08630915A
; Patent No. 6309820
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: HOFFMAN, No. 6309820h
; APPLICANT: KAY, Brian K.
; APPLICANT: FOWLKES, Dana M.
; APPLICANT: MCCONNELL, Stephen J.
; TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
; TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
; TITLE OF INVENTION: USING SAME
; NUMBER OF SEQUENCES: 227
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/630,915A
; FILING DATE: 03-APR-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: MISROCK, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-174
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4091 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-630-915A-33

```

	Query Match	10.4%	Score 29.6;	DB 4;	Length 4091;
	Best Local Similarity	51.5%	Pred. No. 1.7;		
	Matches 68;	Conservative 0;	Mismatches 64;	Indels 0;	Gaps 0;
Qy	10	ctacagctgaccagctgccgctatcgcaatgtgttgaggccctctcgaggattata	69		
Db	995	CGAGAACTGGAGCAGGCCATCCGGGGGCCGATGCCAGGAGGACCTCAGGTGGTTCCGC	1054		
Qy	70	agaaacggaggcgctatggaggccatgaggggggtgaactcacagcaacaggcgccaggg	129		
Db	1055	AGCACCAGTGCCCGCGGATGCCCATGAAC TGCCCGCATTCGAGGAGTGGAAACCAGAC	1114		
Qy	130	ctgcccacgcc	141		
Db	1115	CTCCCGCACACC	1126		

RESULT

```

US-08-202-389-1
: Sequence 1, Application US/08202389
: Patent No. 5536636
: GENERAL INFORMATION:
: APPLICANT: Freeman Jr., Robert M.
: APPLICANT: Plutzky, Jorge
: APPLICANT: Neel, Benjamin G.
: APPLICANT: Rosenberg, Robert D.
: TITLE OF INVENTION: IDENTIFICATION OF NOVEL TYROSINE
: TITLE OF INVENTION: PHOSPHATASES HAVING SH2 DOMAINS
: NUMBER OF SEQUENCES: 54
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
: STREET: Two Militia Drive
: CITY: Lexington
: STATE: MA
: COUNTRY: USA
: ZIP: 02173
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/202,389
: FILING DATE: 28-FEB-1994
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/983,926
: FILING DATE: 01-DEC-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/829,141
: FILING DATE: 31-JAN-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/721,112
: FILING DATE: 26-JUN-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Granahan, Patricia
: REGISTRATION NUMBER: 32,27
: REFERENCE/DOCKET NUMBER: BIH92-05MA
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 861-6240
: TELEFAX: (617) 861-9540
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1747 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: CDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 2..1540
: US-08-202-389-1

```

Query Match

Best Local Similarity 57.0%; Pred. No. 1.8;
Matches 53; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

Qy	96	gaaggggctgaacgctcaacgacaaacaggcgagggcgtgccacgcgccttattttgcctg	155
Db	105	GAGGCGAGCAGATCAGTCGTGCAAGCCAAAGGGCGAGCCCTGGACATTTCTTGTCGGT	164
Qy	156	ctacgaaaggtataaaaagacattgagtgt	188
Db	165	AGAGTCTCAGCCAAACCTGGTGATTTTGTGCTCT	197

RESULT 7

US-09-457-708-1/c

; Sequence 1, Application US/09457708
; Patent No. 6326483

; GENERAL INFORMATION:
 ; APPLICANT: Kwiatkowski, David J.
 ; APPLICANT: Sampson, Julian R.
 ; APPLICANT: Povey, Sue
 ; APPLICANT: van Slegtenhorst, Marjon
 ; APPLICANT: Halley, Dicky
 ; TITLE OF INVENTION: Compositions and Methods Based U
 ; TITLE OF INVENTION: Sclerosis-1 (TSC-1) Gene and Gene Product
 ; NUMBER OF SEQUENCES: 28
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Vinson & Elkins
 ; STREET: 1455 Pennsylvania Avenue, N.W.
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: U.S.
 ; ZIP: 20004-1008
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/457,708
 ; FILING DATE:
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Sanzo, Michael A.
 ; REGISTRATION NUMBER: 36,912
 ; REFERENCE/DOCKET NUMBER: BRI331/42002
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 639-6585
 ; TELEFAX: (202) 639-6604
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 8600 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; US-09-457-708-1

Query Match 10.2%; Score 29; DB 4; Length 8600;
Best Local Similarity 58.8%; Pred. No. 3.8;
Matches 50; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 100 ggcctgaacgtcacagcaaacgagcgaggccctgccccagcccttatttttcgtgtac 159
||| | ||||||| |||. |||| | ||| ||| ||| ||| |||
7750 GGGCAAGGGTACACAGCAGCAGCCTAGGGGCCACTCCTCGACCATAAACTTGTTCTTTTCAC 7691

Qy 160 gaaaagttaaaaaagacattgagtg 184
 ↑ ↑ ↑ ↑ ↑ ↑ ↑ ↑
 Db 7690 CTACAGACAAAAGCTTAATCAAGTG 7666

RESULT

```

US-08-026-138E-17/c
; Sequence 17, Application US/08026138E
; Patent No. 5502166
; GENERAL INFORMATION:
; APPLICANT: Masayoshi MISHINA
; TITLE OF INVENTION: NOVEL PROTEINS AND GENES CODING THE SAME
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nishiohata Residence I-107
; STREET: 5214, Nishiohata-machi
; CITY: Niigata-shi
; STATE: Niigata-ken
; COUNTRY: JAPAN

```

```

; FILING DATE: 26-FEB-1993
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: JP 39563/1992
; FILING DATE: 26-FEB-1992
; APPLICATION NUMBER: JP 173155/1992
; FILING DATE: 30-JUN-1992
; APPLICATION NUMBER: JP 215017/1992
; FILING DATE: 12-AUG-1992
; APPLICATION NUMBER: JP 303878/1992
; FILING DATE: 13-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Hamburg, C.Bruce
; REGISTRATION NUMBER: 22,389
; REFERENCE/DOCKET NUMBER: F-4551
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 986-2340
; TELEFAX: (212) 953-7733
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4446 nucleic acids
; TYPE: nucleic acid
; STRANDEDNESS: double strand
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; ORIGINAL SOURCE:
; ORGANISM: mouse
; TISSUE TYPE: brain
; PUBLICATION INFORMATION:
; AUTHORS: Masayoshi MISHINA
; TITLE: NOVEL PROTEINS AND GENES CODING THE SAME
; RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 1 TO 4446
; US-08-026-138E-6

Query Match 10.0%; Score 28.4; DB 1; Length 4446;
Best Local Similarity 70.4%; Pred. No. 4.5;
Matches 38; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 64 attataaacggaggccctatgagcccatgagggcgctgaacgtcacagca 117
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 401 ATTATCATAGATGAGCCCCCATGATGCCAGGATGGGGGTGAGAGTCTGAGCA 348

RESULT 10
US-09-268-163-9/C
; Sequence 9, Application US/09268163B
; Patent No. 6353091
; GENERAL INFORMATION:
; APPLICANT: Lipscombe, Diane
; TITLE OF INVENTION: HUMAN N-TYPE CALCIUM CHANNEL ISOFORM AND USES THEREOF
; FILE REFERENCE: BL055/7000
; CURRENT APPLICATION NUMBER: US/09/268,163B
; CURRENT FILING DATE: 1999-03-12
; EARLIER APPLICATION NUMBER: US 60/077,901
; EARLIER FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 7011
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..7008
; US-09-268-163-9

Query Match 9.9%; Score 28.2; DB 4; Length 7011;
Best Local Similarity 54.3%; Pred. No. 6.6;
Matches 57; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

Ov 167 taaaaaacacattgaatgatgaatccaccctgaggaacaaatagcccatattaccatattg 226

```


; REFERENCE/DOCKET NUMBER: 7326-006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3771 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: CDNA
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 345..2558
US-08-185-432-3

Query Match 9.5%; Score 27.2; DB 1; Length 3771;
Best Local Similarity 49.3%; Pred. No. 11;
Matches 71; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

Qy 9 tctacagctgaccagctgccgcgtatcgcaatgtgttgaggccctctgaggattat 68
Db 1035 TGTACTGGCTGCTCAGCTGGCGGAATTCGACTTGAGTTGGTGGCCTGTGTGCGGTTA 976

Qy 69 aagaacggaggcctatgaggcccatgaggggcgtgaacgtcacagcaacagggcgag 128
Db 975 GTTTCACCAAGGATACGCGCCTGTGGTACGCCGAATGCTGGCATGGTCCGCTGG 916

Qy 129 gcctgcccacgcccctttatttgc 152
Db 915 GTTGGCGCACGTGGGTGAGATTGC 892

RESULT 14
US-08-185-432-23
; Sequence 23, Application US/08185432
; Patent No. 5750652
; GENERAL INFORMATION:
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Busseau, Isabelle
; APPLICANT: Diederich, Robert J.
; APPLICANT: Xu, Tian
; APPLICANT: Matsuno, Kenji
; TITLE OF INVENTION: DELTEX PROTEINS, NUCLEIC ACIDS, AND
; TITLE OF INVENTION: ANTIBODIES, AND RELATED METHODS AND COMPOSITIONS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/185,432
; FILING DATE: 21-JAN-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 23:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 3771 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: CDNA
US-08-185-432-23

Query Match 9.5%; Score 27.2; DB 1; Length 3771;
Best Local Similarity 49.3%; Pred. No. 11;
Matches 71; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

Qy 9 tctacagctgaccagctgccgcgtatcgcaatgtgttgaggccctctgaggattat 68
Db 2737 TGTACTGGCTGCTCAGCTGGCGGAATTCGACTTGAGTTGGTGGCCTGTGTGCGGTTA 2796

Qy 69 aagaacggaggcctatgaggcccatgaggggcgtgaacgtcacagcaacagggcgag 128
Db 2797 GTTTCACCAAGGATACGCGCCTGTGGTACGCCGAATGCTGGCATGGTCCGCTGG 2856

Qy 129 gcctgcccacgcccctttatttgc 152
Db 2857 GTTGGCGCACGTGGGTGAGATTGC 2880

RESULT 15
US-08-475-844-8
; Sequence 8, Application US/08475844
; Patent No. 5972843
; GENERAL INFORMATION:
; APPLICANT: Lobanenko, Victor V.
; APPLICANT: Neiman, Paul E.
; APPLICANT: Klenova, Elena M.
; APPLICANT: Goodwin, Graham H.
; APPLICANT: Filippova, Galina N.
; APPLICANT: Collins, Steven J.
; TITLE OF INVENTION: CTCF
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: One Market Plaza
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,844
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/261,680
; FILING DATE: 17-JUN-1994
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14538A-11-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-467-9600
; TELEFAX: 415-543-5043
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3810 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA

```
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: human CTCF cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 292..2475
; FEATURE:
; NAME/KEY: exon
; LOCATION: 281..1074
; OTHER INFORMATION: /label= exon2
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1075..1245
; OTHER INFORMATION: /label= exon3
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1246..1379
; OTHER INFORMATION: /label= exon4
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1380..1499
; OTHER INFORMATION: /label= exon5
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1500..1649
; OTHER INFORMATION: /label= exon6
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1650..1810
; OTHER INFORMATION: /label= exon7
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1810..1992
; OTHER INFORMATION: /label= exon8
;
US-08-475-844-8
```

```
Query Match          9.5%; Score 27.2; DB 2; Length 3810;
Best Local Similarity 64.1%; Pred. No. 11;
Matches 41; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 45 gttgaggccctctgaggattataagaacgaggccctatggagggcccatgaggggct 104
    |||||
Db 682 GTACAAGAACCTTCAGGGGCTTATGAAATGAAGTCTTAAGAGGGCCTTGCGAAAGT 741

Qy 105 gaac 108
    ||||
Db 742 GAAC 745
```

Search completed: September 28, 2002, 07:56:00
Job time: 11245 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2002, 06:22:48 ; Search time 4881.77 Seconds
(without alignments)
787.959 Million cell updates/sec

Title: US-09-870-113-9

Perfect score: 285

Sequence: 1 atcagaggtctacagcctga.....tgaacctgcggaaggtga 285

Scoring table:

IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	280	98.2	747	10	BI827322 603077989
2	277.6	97.4	888	9	AL530804 AL530804
3	265.8	93.3	480	10	BI045863
4	257.4	90.3	573	10	BI541703
5	246.8	86.6	598	10	BE913718
6	246.8	86.6	645	10	BG964218
7	246.8	86.6	1131	10	BG295496
8	245.2	85.0	560	10	BI794616
9	244.6	85.8	574	10	BI185580
10	234	82.1	501	9	AW211366
11	224.4	78.7	701	10	BI831943
12	223.4	78.4	336	9	AW326482
13	223.2	78.3	817	10	BI103329
14	195.4	68.6	546	10	BM488747
15	191.8	67.3	855	10	BI412175
16	191.6	67.2	453	10	BF916224
17	188.4	66.1	967	10	BI739710

18	185.6	65.1	397	9	BE012485
19	182.2	63.9	697	10	BI854643
20	168	58.9	722	10	BI876051
21	166.2	58.3	622	10	BJ039504
22	166	58.2	681	9	AV704087
23	156.8	55.0	421	10	BI448355
24	149.8	52.6	637	9	AL654963
25	143.8	50.5	417	9	AA104365
26	142.6	50.0	546	9	BE014800
27	140.8	49.4	561	9	AI942584
28	139.4	48.9	678	9	BB654057
29	139.2	48.8	1295	11	AK015770
30	137.8	48.4	916	10	BI757180
31	129.6	45.5	567	10	BJ035232
32	129.6	45.5	612	10	BJ012200
33	128.2	45.0	512	9	AW210205
34	128	44.9	684	9	AI133696
35	128	44.9	750	10	BI831750
36	126.4	44.4	488	9	AI310713
37	126	44.2	309	9	AI630395
38	125.4	44.0	729	10	BI460337
39	116	40.7	803	10	BG777404
40	115.2	40.4	1209	11	AK006155
41	113.8	39.9	331	10	BI449976
42	109.8	38.5	666	9	BB644566
43	107.6	37.8	593	10	BG895490
44	105.2	36.9	483	9	AA061624
45	103.4	36.3	350	12	AZ693917

ALIGNMENTS

BI827322 747 bp mRNA linear EST 04-OCT-2001
603077989F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5169694 5',
mRNA sequence.

BI827322 GI:15938872

EST.

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM1422 row: a column: 23

High quality sequence stop: 743.

Location/Qualifiers

1. .747

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5169694"

/clone_lib="NIH_MGC_119"

/tissue_type="medulla"

/lab_host="DH10B"

/note="Organ: brain; Vector: pCMV-SPORT6; Site:1: NotI;

Site:2: EcoRV (destroyed); RNA source normal medulla from

anonymous male age 27. Library is oligo-dT primed and

directionally cloned (EcoRV site is destroyed upon

cloning). Average insert size 1.3 kb, insert size range

0.9-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 013. Note: this is a NIH_MGC Library."

BASE COUNT 188 a 206 c 201 g 151 t 1 others
ORIGIN

Query Match 98.2%; Score 280; DB 10; Length 747;
Best Local Similarity 100.0%; Pred. No. 3.1e-76;
Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgcagagctctacagccctgagccagctgcccgcctatcgcaatgtgttgagggccctctgg 60
|||||
Db 84 ATGCAGAGTCTACAGCCTGACCCAGCTGCCGCTATGCGAATGTGTGGAGGCCCTCTGG 143
|||||
QY 61 aggtattataagaacagggagggccctatggagggccatgaggggctgaaagctcacagcaaca 120
|||||
Db 144 AGGATTATAAGAAACAGAGGGCCCTATGGAGGCCATGAGGGGCTGAACGTCACAGCAACA 203
|||||
QY 121 ggcgagggcctgccagcagcccttatttgcctgctctacgaaagttaaaagacattg 180
|||||
Db 204 GCGCAGGGCCCTGCCACGCGCCCTTTATTTCCTGCTACGAAAGTTAAAAAGACATTG 263
|||||
QY 181 agtgatgtaataccaccctgggggcaatagccatattgccaatggcgccgggtgtgtg 240
|||||
Db 264 AGTGATGTAATCCACCTGGGGGCAATAGCCATATGTCATGTCGCGCGGGGTGTGTG 323
|||||
QY 241 gcaacattactctatgcatgagccatgaacctgcgggaag 280
|||||
Db 324 GCAACATTACTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 363
|||||

RESULT 2

AL530804
LOCUS AL530804 LTI_NFL001_NBC4 888 bp mRNA linear EST 13-FEB-2001
DEFINITION prime, mRNA sequence.

ACCESSION AL530804
VERSION AL530804.1 GI:12794297
KEYWORDS EST.

SOURCE

human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 888)

Li.W.B., Gruber,C., Jessee,J. and Polayes,D.

Full-length cDNA libraries and normalization

Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage

Bp 191 91006 EVRY cedex - France

Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr.

Location/Qualifiers

1. .888

FEATURES

source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSDD008TH17"
/clone_lib="LTI_NFL001_NBC4"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"
BASE COUNT 168 a 244 c 322 g 139 t 15 others

ORIGIN

Query Match 97.4%; Score 277.6; DB 9; Length 888;
Best Local Similarity 99.3%; Pred. No. 1.8e-75;
Matches 278; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 atgcagagctctacagccctgagccagctgcccgcctatcgcaatgtgttgagggccctctgg 60
|||||
Db 417 ATGCAGAGTCTACAGCCTGACCCAGCTGCCGCTATGCGAATGTGTGGAGGCCCTCTGG 476
|||||
QY 61 aggtattataagaacagggagggccctatggagggccatgaggggctgaaagctcacagcaaca 120
|||||
Db 477 AGGATTATAAGAAACAGAGGGCCCTATGGAGGCCCATGAGGGGCTGAACGTCACAGCAACA 536
|||||
QY 121 ggcgagggcctgccagcagcccttatttgcctgctctacgaaagttaaaagacattg 180
|||||
Db 537 GCGCAGGGCCCTGCCACGCGCCCTTTATTTCCTGCTCAGCAAAAGTTAAAAAGACATTG 596
|||||
QY 181 agtgatgtaataccaccctgggggcaatagccatattgccaatggcgccgggtgtgtg 240
|||||
Db 597 AGTGATGTTATCCACCTGGGGGCAATAGCCATATGTCATGTCGCGCGGGGTGTGTG 656
|||||
QY 241 gcaacattactctatgcatgagccatgaacctgcgggaag 280
|||||
Db 657 GCAACATTACTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 696
|||||

RESULT 3

BI045863
LOCUS BI045863 MR3-FN0209-300101-004-h10 FN0209 Homo sapiens cDNA, mRNA sequence.
DEFINITION

ACCESSION BI045863

VERSION BI045863.1 GI:14452485

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 480)

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,

Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,

Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hate

M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and

Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR3&t2=MR3-FN0209-

300101-004-h10&t3=2001-01-30&t4=1)

Seq primer: puc 18 forward

High quality sequence stop: 480.

Location/Qualifiers

1. .480

FEATURES

source

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="FN0209"

/dev_stage="Adult"

/note="Organ: prostate_normal; Vector: puc18; Site_1: Smal

; Site_2: Smal; A mini-library was made by cloning

products derived from ORESTES PCR (U.S. Letters Patent

application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

116 a 126 c 146 g 92 t

BASE COUNT 116 a 126 c 146 g 92 t

ORIGIN

Query Match 93.3%; Score 265.8; DB 10; Length 480;
Best Local Similarity 98.9%; Pred. No. 7.2e-72;
Matches 278; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 atgcagagctctacagcctgaccagctgcccgcctatcgcaatgtgttgaggccctctgg 60
|||||
Db 74 ATGCAGAGCTACAGCTACCCAGCTGCCGCTATCGCAATGTGTGAGGCCCTCTGG 133
|||||

QY 61 aggtattataagaaacgagggcctatgagcccatgagggctgaacctcacagcaaca 120
|||||
Db 134 AGGATTATTAAGAACGAGGGCCCTATGGAGCCCATGAGGGGCTGAACGTGCACAGCAACA 193
|||||

QY 121 ggcgcaggcctgcccacgcccctttatttgcctgctacgaaagttaaaaaagacattg 180
|||||
Db 194 GCGCAGGGCCCTGCCACACGCGCTTTATTTTTCCTGCTACGAAAGTTAAAAAGACATTG 253
|||||

QY 181 agtgatgtaataccacct-gggggcaatagcgaatattgcaatgtgtgcggcggtgtgt 239
|||||

Db 254 AGTGATGTAAATCCACCTCGGGGGCAATAGCCATATGTGCAATGTGTGCGCGGGTGTGT 313
|||||

QY 240 ggcacattacttcatgatgcagccatgaacctgcggaag 280
|||||

Db 314 GGCACCATTTACTTCATGATGACGCCATGAACCTGCGGAAG 354
|||||

RESULT 4
BI541703
LOCUS BI541703 573 bp mRNA linear EST 30-AUG-2001
DEFINITION 455441 MARC IBOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION BI541703
VERSION BI541703.1 GI:15382815
KEYWORDS EST.
SOURCE Bos taurus
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.
Smith, T. P. L., Grosse, W. M., Freking, B. A., Roberts, A. J., Stone, R. T., Casas, E., Wray, J. E., White, J., Cho, J., Fahrnerkrug, S. C., Bennett, G. L., Heaton, M. P., Laegreid, W. W., Rohrer, G. A., Chitko-McKown, C. G., Perlee, G., Holt, I., Karamycheva, S., Liang, F., Quackenbush, J. and Keefe, J. W.
1 (bases 1 to 573)
REFERENCE Smith, T. P. L., Freking, B. A., Roberts, A. J., Stone, R. T., Casas, E., Wray, J. E., White, J., Cho, J., Fahrnerkrug, S. C., Bennett, G. L., Heaton, M. P., Laegreid, W. W., Rohrer, G. A., Chitko-McKown, C. G., Perlee, G., Holt, I., Karamycheva, S., Liang, F., Quackenbush, J. and Keefe, J. W.
AUTHORS
TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle
JOURNAL Genome Res. 11 (4), 626-630 (2001)
MEDLINE 21180013
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smithemall.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTTCCAGTCAGCAGC
Plate: 121 row: O column: 24
Seq primer: ATTAGGTGACACTATAG.
Location/Qualifiers
1..573
/organism="Bos taurus"
/db_xref="taxon:9913"

/clone_lib="MARC IBOV"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI; Library made from pooled tissue from lymph node, ovary, fat, hypothalamus, and pituitary."

121 a 166 c 176 g 110 t

BASE COUNT 121 a 166 c 176 g 110 t

ORIGIN

Query Match 90.3%; Score 257.4; DB 10; Length 573;
Best Local Similarity 94.3%; Pred. No. 3.1e-69;
Matches 267; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 atgcagagctctacagcctgaccagctgcccgcctatcgcaatgtgttgaggccctctgg 60
|||||
Db 174 ATGCAGAGCTACAGCGGATCCAGCGCCGCTATCGCAATGTGTGAGGCCCTCTGG 233
|||||

QY 61 aggtattataagaaacgagggcctatgagcccatgagggctgaacctcacagcaaca 120
|||||
Db 234 AGGATTATTAAGAACGAGGGCCCTGTGGAGGCCCATCGAGGCCGTGAACGTCACAGCAACA 293
|||||

QY 121 ggcgcaggcctgcccacgcccctttatttgcctgctacgaaagttaaaaaagacattg 180
|||||
Db 294 GCGCAGGGCCCTGCCACACGCGCTTATTTTTCCTGCTATGAAAGTTAAAAAGACATTG 353
|||||

QY 181 agtgatgtaataccacctgggggcaatagcgaatattgcaatgtgtgcggcggtgtgtg 240
|||||

Db 354 AGTGATGTAAATCCACCTCGGGGGCAATAGCCATATGTGCAATGTGTGCGCGGTGTGTG 413
|||||

QY 241 gcaacattacttcatgatgcagccatgaacctgcggaaggt 283
|||||

Db 414 GCAACATTACTTCATGATGCTGCCATGAATCCAGTGGAAAGACT 456
|||||

RESULT 5
BE913718
LOCUS BE913718 598 bp mRNA linear EST 29-SEP-2000
DEFINITION 601669072F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:3969256 5', mRNA sequence.
ACCESSION BE913718
VERSION BE913718.1 GI:10411618
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 598)
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9146 row: O column: 17
High quality sequence stop: 593.
Location/Qualifiers
1..598
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:3969256"
/clone_lib="NCI_CGAP_Mam1"
/tissue_type="tumor, biopsy sample"
/dev_stage="10 months, virgin"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;

Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

BASE COUNT 141 a 175 c 172 g 110 t

ORIGIN

Query Match 86.6%; Score 246.8; DB 10; Length 598;
Best Local Similarity 92.2%; Pred. No. 6.2e-66;
Matches 260; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 1 atgcagagctacagcctgaccagctgcccgcctatcgcaatgtgttgaggccctctgg 60

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 104 ATGCAGAGCTACAGCCTGACCCAGCCGCCGCTATCGAAGCTTTGGAGGCTCTCTGG 163

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Qy 61 aggattataaagagggcctatggagccatgaagggtgaacgtcacagcaaca 120

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 164 AGAATCATGAGACAGAGGGCCCTGTGGAGGCCCAATGCGGGGCTCAAGCTCACAGCAACA 223

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Qy 121 ggcgcaggcctgcccagcgcctttatttgcctgctacgaaaagttaaaaagacattg 180

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 224 GCGCGGGCCTGCCACGCCCTCTATTTGCTGCTACGAAAGTTAAAAGACATTG 283

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Qy 181 agtcatgtaataccaccctgggggcaatagccatatgtccaatggtgcggccgggtgtgtg 240

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 284 AGTGACGTAATCCACCCAGGGGCAATAGCCATATTTGCCAATGTGCAAGCGGGATGTGTG 343

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Qy 241 gcaacattactcatgacgcacatgaaccctgcggaagc 282

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 344 GCGACATTACTTCATGATGCAGCCATGAATCCAGCGGAAGTC 385

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

RESULT 6

BC964218 BC964218 645 bp mRNA linear EST 12-JUN-2001

LOCUS 602828996F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4983815 5',

DEFINITION mRNA sequence.

ACCESSION BC964218

VERSION BC964218.1 GI:14351855

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

TITLE 1 (bases 1 to 645)

JOURNAL NIH-MGC http://mgc.nci.nih.gov/

COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM10988 row: p column: 24

High quality sequence start: 2

High quality sequence stop: 632.

Location/Qualifiers

1. 645

/organism="Mus musculus"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="IMAGE:4983815"

/clone_lib="NCI_CGAP_Co24"

/lab_host="DH10B (T1 phage-resistant)"

/note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;

Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.

Average insert size 1.6 kb. Constructed by Life

Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 147 a 189 c 193 g 116 t

ORIGIN

Query Match 86.6%; Score 246.8; DB 10; Length 645;

Best Local Similarity 92.2%; Pred. No. 6.3e-66;

Matches 260; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 1 atgcagagctacagcctgaccagctgcccgcctatcgcaatgtgttgaggccctctgg 60

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 173 ATGCAGAGCTACAGCCTGACCCAGCCGCCGCTATCGAAGCTTTGGAGGCTCTCTGG 232

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Qy 61 aggattataaagagggcctatggagcccatgaggggctgaacgtcacagcaaca 120

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 233 AGAATCATGAGACAGAGGGCCCTGTGGAGGCCCAATGCGGGGCTCAAGCTCACAGCAACA 292

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Qy 121 ggcgcaggcctgcccagcgcctttatttgcctgctacgaaaagttaaaaagacattg 180

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 293 GCGCGGGCCTGCCACGCCCTCTATTTGCTGCTACGAAAGTTAAAAGACATTG 352

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Qy 181 agtcatgtaataccaccctgggggcaatagccatatgtccaatggtgcggccgggtgtgtg 240

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 353 AGTGACGTAATCCACCCAGGGGCAATAGCCATATTTGCCAATGTGCAAGCGGGATGTGTG 412

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Qy 241 gcaacattactcatgacgcacatgaaccctgcggaagc 282

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 413 GCGACATTACTTCATGATGCAGCCATGAATCCAGCGGAAGTC 454

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

RESULT 7

BC295496 BC295496 1131 bp mRNA linear EST 21-FEB-2001

LOCUS 602392826F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:4504638 5',

DEFINITION mRNA sequence.

ACCESSION BC295496

VERSION BC295496.1 GI:13057189

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

TITLE NIH-MGC http://mgc.nci.nih.gov/

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: The Cepko Laboratory

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Clone distribution: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM10377 row: c column: 07

High quality sequence stop: 753.

Location/Qualifiers

1..1131

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone="IMAGE:4504638"

/clone_lib="NIH_MGC_94"

/tissue_type="retina"

/lab_host="DH10B (phage-resistant)"

/note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;

Site_2: SalI; Cloned unidirectionally; oligo-dt primed.

Average insert size 3.3 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH_MGC Library."

BASE COUNT 261 a 335 c 375 g 160 t

ORIGIN

Query Match 86.6%; Score 246.8; DB 10; Length 1131;

Best Local Similarity 92.2%; Pred. No. 7.3e-66;
Matches 260; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 1 atgcagagctacagcctgaccagctgccgctatcgaatgtgttgagccctctgg 60
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 261 ATGCAGAGCTACAGCTGACCCAGCGCCGCTATCGAAGCTGTGGAGGCTCTCTGG 320
QY 61 aggattataagaacagagggccctatggagggccatgaggggctgaacgtcacagaaca 120
Db 321 AGAATCATGAGACAGAGGCGCTGTGGAGGCCATCGCGGGGCTGAACGTCACAGCAACA 380
QY 121 ggcgagggcctgccagcagcccttatttctgctctacgaaagttaaaaaagacattg 180
Db 381 GCGCGGGGCGCTGCCACCGCCCTCTATTGTGCTCTACGAAAGCTTAAAAAGACATTG 440
QY 181 agtgtgttaataccacctgggggcaatgacccatattgccaatggtgcgcggtgtgtg 240
Db 441 AGTGAGTAATCCAGCGGGGCAATAGCCATATTGCCAATGGTGACCGCGAATGTGTG 500
QY 241 gcaacattactctatgcatgacccatgaacctgcggaagc 282
Db 501 GCGAATTACTTTCATGATGACGCCATGAATCCAGCGGAAGTC 542

RESULT 8
BI794616
LOCUS
DEFINITION
ic92e02.y1 Melton Normalized Mixed Mouse Pancreas 1 NI-MMS1 Mus
musculus cDNA 5' similar to TR:Q23125 Q23125 W02B12.9 PROTEIN. [1]
; mRNA sequence.
ACCESSION
BI794616
VERSION
BI794616.1 GI:15822341
KEYWORDS
EST.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 560)
Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
Lemishka,I., Searce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
,M., Gibbons,M., Cole,R., Cole,R., Tsagarishvili,R., Williams,I.,
, Jackson,Y. and Bowers,Y.
Endocrine Pancreas Consortium
Unpublished (2000)
Other ESTs: ic92e02.x1
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Douglas Melton DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Juliana Brown
(brown@fas.harvard.edu)
Seg primer: 40RP from Gibco
High quality sequence stop: 430.
Location/Qualifiers
1..560
/organism="Mus musculus"
/strain="ICR"
/db_xref="taxon:10090"
/clone_lib="Melton Normalized Mixed Mouse Pancreas 1
NI-MMS1"
/sex="Both for embryonic & newborn, male for adult and
adult islet"
/dev_stage="Embryonic day 10.5, E12.5, E16.5, newborn,
adult, mixed"
/lab_host="DH10B"

/note="Vector: pSPORT1; Site_1: Not I; Site_2: Sal I; Five
libraries representing E10.5/12.5 pancreatic bud, E16.5
pancreas, newborn pancreas, adult pancreas, and adult
islets of Langerhans were separately constructed using
SuperScript Plasmid Library kit (Life Technologies). cDNA
was made by oligo-dT priming and size-selected by column
fractionation. Libraries were amplified once on solid
support and plasmid DNA from each library was prepared
and mixed in equal amounts. The mixed library DNA was
normalized by method #4 from Bonaldo, Lennon, and Soares
1996 Genome Research 6:791-806; 0.5 microgram
single-stranded mixed library plasmid DNA was mixed with
5 micrograms PCR product representing mixed library
inserts and hybridized to an EcoT 6. Single-stranded
(unhybridized) plasmids were isolated by hydroxyapatite
chromatography and used to make this library."
BASE COUNT 134 a 158 c 163 g 105 t
ORIGIN

Query Match 86.0%; Score 245.2; DB 10; Length 560;
Best Local Similarity 91.8%; Pred. No. 1.9e-65;
Matches 259; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 1 atgcagagctacagcctgaccagctgccgctatcgaatgtgttgagccctctgg 60
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 84 ATGCAGAGCTACAGCTGACCCAGCGCCGCTATCGAAGCTGTGGAGGCTCTCTGG 143
QY 61 aggattataagaacagagggccctatggagggccatgaggggctgaacgtcacagaaca 120
Db 144 AGAATCATGAGACAGAGGCGCTGTGGAGGCCATCGCGGGGCTGAACGTCACAGTAACA 203
QY 121 ggcgagggcctgccagcagcccttatttctgctctacgaaagttaaaaaagacattg 180
Db 204 GCGCGGGGCGCTGCCACCGCCCTCTATTGTGCTCTAGAAAAAGCTTAAAAAGACATTG 263
QY 181 agtgtgttaataccacctgggggcaatgacccatattgccaatggtgcgcggtgtgtg 240
Db 264 AGTGAGTAATCCAGCGGGGCAATAGCCATATTGCCAATGGTGACCGCGAATGTGTG 323
QY 241 gcaacattactctatgcatgacccatgaacctgcggaagc 282
Db 324 GCGAATTACTTTCATGATGACGCCATGAATCCAGCGGAAGTC 365

RESULT 9
BI185580
LOCUS
DEFINITION
UNL-P-FN-cy-f-02-0-UNL.s1 UNL-P-FN Sus scrofa cDNA clone
UNL-P-FN-cy-f-02-0-UNL 3', mRNA sequence.
ACCESSION
BI185580
VERSION
BI185580.1 GI:14659989
KEYWORDS
EST.
SOURCE
pig.
ORGANISM
Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE
1 (bases 1 to 574)
Caetano,A.R., Johnson,R.K. and Pomp,D.
Generation and sequence characterization of a normalized cDNA
library from swine ovarian follicles
Unpublished (2001)
Contact: Pomp, D
Department of Animal Science
University of Nebraska, Lincoln
Lincoln, NE 68583-0908, USA
Tel: 402 472 6416
Fax: 402 472 6362
Email: dpomp@unl.edu
Oligo-dT track not found, Not I site shown in beginning of sequence
is likely internal to the message.
Seq primer: M13 -29
POLYA=No.

```
FEATURES
source
Location/Qualifiers
1. .574
/organism="Sus scrofa"
/strain="University of Nebraska, Lincoln Swine Selection
Lines"
/db_xref="taxon:9823"
/clone="UNL-P-FN-cy-f-02-0-UNL"
/clone_lib="UNL-P-FN"
/dev_stage="ADULT"
/lab_host="DH10B (Life Technologies)"
/notes="vector: p773D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; The UNL-P-FN
library is a normalized library representing porcine
ovarian follicles, ranging between 2.0 to 10.0 mm in
diameter, collected during 7 days of the follicular phase
of the pig estrous cycle. This library was derived from
the library UNL-P-F2. The tag is a string of 5-6
nucleotides present between the Not I site and the
oligo-dT track. The library was constructed as described
by Bonaldo, Lennon and Soares, Genome Research 6: 791-806
, 1996.
TAG_SEQ=None found"
BASE COUNT 132 a 171 c 165 g 104 t 2 others
ORIGIN
Query Match 85.8%; Score 244.6; DB 10; Length 574;
Best Local Similarity 94.3%; Pred. No. 2.9e-65;
Matches 264; Conservative 0; Mismatches 15; Indels 1; Gaps 1;
Qy 1 atgcagagctcacagcctgacccagctgcccgtctatgcgaatgtgttgaggccctctgg 60
Db 135 ATGCAGAGCCTGCAGCCTGATCCCGCCGCCCTATGCAATGTGTGGAGGCCCTCTGG 194
Qy 61 aggtattataagaacgaggggcctatgaggcccatgaggggctgaaactcacagcaaca 120
Db 195 AGGATTATATAAGAACGAGGGGCTGTGGAGGCCATGCGGGGCTGAACGTCACAGCAACA 254
Qy 121 ggcagaggcctgccagcgccttatttttcctgctacgaaaaagttaaaaaagacattg 180
Db 255 GCGCAGGGCCTGCCAGCGCCTCTATTTTCCTGCTATGAAAGTTAAAAAAGACATTG 314
Qy 181 agtgatgtaatccacctgggggcaatagccatatgtccaatgtgctgaggggtgtgtg 240
Db 315 AGTGATGTAATCCACCTTGGGGCAATAGCCATATAGCCATGTGTGGCGCGGTGGCT- 373
Qy 241 gcaacattactctatgctgagccatgaacctgcggaag 280
Db 374 ANAACATTACTTTCATGTCAGCCATGAATCCAGAGAAG 413
RESULT 10
AW211366 501 bp mRNA linear EST 03-DEC-1999
LOCUS uc080b07.y1 NCI_CGAP_Mam3 Mus musculus cDNA clone IMAGE:2648821 5'
DEFINITION similar to WP:W02B12.9 CE03769 MITOCHONDRIAL RNA SPLICING MSR4 LIKE
PROTEIN ; mRNA sequence.
ACCESSION AW211366
VERSION AW211366.1 GI:6517318
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 501)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
CONTACT: Robert Strausberg, Ph.D.
Email: cgaps-f@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
```

```
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
MGI:1029273
Seq primer: -40RP from Gibco
High quality sequence stop: 427.
location/Qualifiers
1. .501
/organism="Mus musculus"
/strain="129 - C57/B6 - FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:2648821"
/clone_lib="NCI_CGAP_Mam3"
/tissue_type="tumor, gross tissue"
/dev_stage="10 months"
/lab_host="DH10B"
/notes="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
Reference for transgenic model: Xu et al., Nature Genetics
22, 37-43 (1999)."
BASE COUNT 120 a 139 c 144 g 98 t
ORIGIN
Query Match 82.1%; Score 234; DB 9; Length 501;
Best Local Similarity 92.5%; Pred. No. 5.6e-62;
Matches 246; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
Qy 1 atgcagagctcacagcctgacccagctgcccgtctatgcgaatgtgttgaggccctctgg 60
Db 53 ATGCAGAGCCTTACAGCCTGACCCAGCGCGCTATCGGAACGTTGTGGAGGCTCTCTGG 112
Qy 61 aggtattataagaacgaggggcctatgagcccatgaggggctgaaactcacagcaaca 120
Db 113 AGAATCATGATGAGGACAGAGGGCCTGTGGAGCCCATGCGGGGCTGAACGTCACAGCAACA 172
Qy 121 ggcagaggcctgccagcgccttatttttcctgctacgaaaaagttaaaaaagacattg 180
Db 173 GCGCGGGGCTGCCAGCGCCTCTATTTTCCTGCTACGAAAGTTAAAAAAGACATTG 232
Qy 181 agtgatgtaatccacctgggggcaatagccatatgtccaatgtgctgaggggtgtgtg 240
Db 233 AGTGACGTAATCCACCCAGGGGCAATAGCCATATGCCAATGTGCGCGGATGTGTG 292
Qy 241 gcaacattactctatgctgagccat 266
Db 293 CGGACATTACTTTCATGATGAGCCCT 318
RESULT 11
BI831943 701 bp mRNA linear EST 04-OCT-2001
LOCUS 603076873F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5168539 5',
DEFINITION mRNA sequence.
ACCESSION BI831943
VERSION BI831943
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 701)
NTH-MGC http://mgc.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Unpublished (1999)
CONTACT: Robert Strausberg, Ph.D.
Email: cgaps-f@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
```

CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L14W1419 row: a column: 20
 High quality sequence start: 29
 High quality sequence stop: 699.
 Location/Qualifiers
 1..701
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5168539"
 /clone_lib="NIH_MGC_119"
 /tissue_type="medulla"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: EcoRV (destroyed); RNA source normal medulla from
 anonymous male age 27. Library is oligo-dr primed and
 directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.3 kb, insert size range
 0.9-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 013. Note:
 this is a NIH_MGC Library."
 BASE COUNT 176 a 167 c 208 g 150 t
 ORIGIN

Query Match 78.7%; Score 224.4; DB 10; Length 701;
 Best Local Similarity 99.6%; Pred. No. 5.9e-59;
 Matches 225; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 atgcagagctctacagcctgaccagctccgcctatcgcaatgtgttgaggccctctgg 60
 Db 302 ATGCAGAGCTCTACAGCTACACCGCTGCGCCCTATGCCAATGTGTGAGGCCCTCTGG 361
 QY 61 agattataagaacgagggccctatggagggccatgaggggctgaagcgtcacagcaaca 120
 Db 362 AGGATTATAAGAACGAGGGCCCTATGGAGGCCCATGAGGGGCTCAACGTCACAGCAACA 421
 QY 121 ggcgagggcctgcccagcccttatttggcctctcagaaaagttaaaagacattg 180
 Db 422 GCGCAGGGCCCTGCCACGCCCCCTTTATTTCCTGCTACGAAAGTTAAAAAGACATTG 481
 QY 181 agtgatgaatccacctgggggccaatagccatatggcaatggcg 226
 Db 482 AGTGATGTAATCCACCCCTGGGGGCAATAGCCATATTGCCATGCTG 527

RESULT 12
 AW326482
 LOCUS AW326482 336 bp mRNA linear EST 25-APR-2001
 DEFINITION 19106 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.
 ACCESSION AW326482
 VERSION AW326482.1 GI:6762403
 KEYWORDS EST.
 SOURCE cow.
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.
 1 (bases 1 to 336)
 Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
 Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett
 ,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.,
 Perlea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and
 Keefe,J.W.
 TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA
 libraries and construction of a gene index for cattle
 JOURNAL Genome Res. 11 (4), 626-630 (2001)
 MEDLINE 21180013

COMMENT

Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel.: 402 762 4366
 Fax: 402 762 4390
 Email: smithemail.marc.usda.gov
 Single pass sequencing. Bases called and trimmed with phred
 v0.980904.e. Vector identified by cross_match with the -minscore 20
 and -mismatch 12 options.

PCR Primers
 FORWARD: AGGAACAGCTATGACCAT
 BACKWARD: GTTTCCAGTCAGCAGC
 Plate: 8 row: F column: 16
 Seq primer: ATTAGTGACACTATAG.

FEATURES

Location/Qualifiers
 1..336
 /organism="Bos taurus"
 /db_xref="taxon:9913"
 /clone_lib="MARC 1BOV"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
 Library made from pooled tissue from lymph node, ovary,
 fat, hypothalamus, and pituitary."

BASE COUNT

72 a 96 c 107 g 61 t

ORIGIN

Query Match 78.4%; Score 223.4; DB 9; Length 336;
 Best Local Similarity 95.4%; Pred. No. 9.8e-59;
 Matches 230; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
 QY 1 atgcagagctctacagcctgaccagctccgcctatcgcaatgtgttgaggccctctgg 60
 Db 96 ATGCAGAGCTCTACAGCGGATCCAGCCGCCCTATGCCAATGTGTGAGGCCCTCTG 155
 QY 61 agattataagaacgagggccctatggagggccatgaggggctgaagcgtcacagcaaca 120
 Db 156 AGGATTATAAGAACGAGGGCCCTGTGGAGGCCATGCGAGGCTCAACGTCACAGCAACA 215
 QY 121 ggcgagggcctgcccagcccttatttggcctctcagaaaagttaaaagacattg 180
 Db 216 GCGCAGGGCCCTGCCACGCCCCCTCTATTTCCTGCTATGAAAGTTAAAAAGACATTG 275
 QY 181 agtgatgaatccacctgggggccaatagccatatggcaatggcgcgcggtgtgtg 240
 Db 276 AGTGATGTAATCCACCCCTGGGGGCAATAGCCATATTGCCATGTCGGCTGGTGTG 335
 QY 241 g 241
 Db 336 G 336

RESULT 13

BI103329

LOCUS

BI103329

DEFINITION

60288838F1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:5043992

TITLE

5', mRNA sequence.

ACCESSION

BI103329

VERSION

BI103329.1 GI:14554222

KEYWORDS

EST.

SOURCE

house mouse.

ORGANISM

Mus musculus

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS

1 (bases 1 to 817)

TITLE

NIH-MGC <http://mgs.nci.nih.gov/>.

JOURNAL

National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT

Unpublished (1999)

CONTACT

Contact: Robert Strausberg, Ph.D.

EMAIL

Email: cgapbs-remail.nih.gov

TISSUE

Tissue Procurement: Jeffrey E. Green, M.D.

CDNA

CDNA Library Preparation: Life Technologies, Inc.

Bonaldo, Ph.D.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLM11294 row: j column: 20
High quality sequence start: 7
High quality sequence stop: 809.
FEATURES
Location/Qualifiers
1..855
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:5120755"
/clone_lib="NCI CGAP Lu33"
/tissue_type="pooled lung tumors"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pT73b-pac (Pharmacia) with a
modified polylinker; Site_1: NotI; Site_2: EcoRI; 1st
strand cDNA was prepared from mRNA obtained from pooled
lung tumors with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGAGCGCGCCCTCTGTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 209 a 233 c 241 g 172 t
ORIGIN

source

Query Match 67.3%; Score 191.8; DB 10; Length 855;
Best Local Similarity 90.6%; Pred. No. 8.4e-49;
Matches 260; Conservative 0; Mismatches 22; Indels 5; Gaps 5;
Qy 1 atgcagagtctacagcctgacccag-ctgccgctatcgcaatgtgtgagggccctctg 59
Db 86 ATGCAGAGCCTACAGCCTGACCCAGACCGCCGCTATCGGAACGTTGTGGAGGCTCTCTG 145
Qy 60 gaggtattataagaacggaggccctatgagggcccatgagggg-gctgaacgtcacagcaa 118
Db 146 GAGATCATGAGGAGAGAGGGGCTGTGGAGGCCCATGCGGGGTGTGAACGTCACAGCAA 205
Qy 119 caggcgagggcc-tgcccacgcccct-ttattttgcctgctacgaaagttaaaaaagac 176
Db 206 CAGGCGGGGGCGCTGCCACGCCCTCGTATTTTCCTGCTACGAAAGTTAAAAAGAC 265
Qy 177 a-ttgagtgtatgaatccaccctggggcaatagccatattgccaatggtcgccgggt 235
Db 266 AGTTGAGTGACGTAATCCACCCAGGGGGCAATAGCCATATTGCCAATGTCGACCGGGAT 325
Qy 236 gtgtggcaacattactcatgatcagccatgaaacccctcggaagggc 282
Db 326 GTGTGGGACATTACTTCATGATGAGCCATGAATCCAGCGGAAGTC 372

Search completed: September 28, 2002, 06:22:49
Job time: 9190 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2002, 07:58:26 ; Search time 5287.54 Seconds
(without alignments)
1567.252 Million cell updates/sec

Title: US-09-870-113-11

Perfect score: 396

Sequence: 1 atcagagtgctacagcctga.....ttacctgtttaccactag 396

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vl.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg.hum.*

31: em.htg.inv.*

32: em.htg.Other.*

33: em.htgo.inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description

1	280	70.7	1322	6	AX061229	Sequence
2	280	70.7	1448	9	AF327402	Homo sapi
3	280	70.7	1448	17	HS303077	Homo sapi
4	273	68.9	401	6	AX071440	Sequence
5	225	56.8	1889	9	AF327403	Homo sapi
6	225	56.8	1889	17	HS303078	Homo sapi
7	225	56.8	123160	9	AL353719	Human DNA
8	202.6	51.2	221062	2	AC096351	Rattus no
9	139.2	35.2	4112	10	AF288621	Mus muscu
10	128	32.3	1429	9	AF223466	Homo sapi
11	125.2	31.6	163205	2	AC099368	Rattus no
12	109.8	27.7	5869	10	AF361699	Mus muscu
13	109.8	27.7	191702	2	AC108878	Mus muscu
14	105	26.5	108765	2	AF216674	Homo sapi
15	104.8	26.5	483	6	AX306258	Sequence
16	103.4	26.1	858	9	BC015013	Homo sapi
17	103.4	26.1	3415	9	AY032628	Homo sapi
18	103.4	26.1	174445	9	AC051642	Homo sapi
19	103.4	26.1	182504	2	AC022597	Homo sapi
20	60.6	15.3	30891	2	AC100095	Mus muscu
21	50	12.6	1072	10	BC023172	Mus muscu
22	49.6	12.5	34305	3	CEW02812	Gaenorhadi
23	42.2	10.7	3037	3	AF060268	Drosophi
24	40.6	10.3	2878	3	AF217402	Drosophi
25	36.2	9.1	1243	3	OGU45997	Onchocera
26	36.2	9.1	1281	3	OVU45998	Onchocera
27	36	9.1	111464	8	F10K1	Sequence
28	35.6	9.0	180397	2	AP003615	Oryza sat
29	35.2	8.9	90935	2	AP003917	Oryza sat
30	35.2	8.9	169672	2	AC103286	Rattus no
31	35	8.8	5747	9	AF108460	Homo sapi
32	35	8.8	6334	9	AF108461	Homo sapi
33	34.6	8.7	223538	2	AC087150	Mus muscu
34	34.4	8.7	1517	9	AF125531	Homo sapi
35	34.4	8.7	205879	2	AL669956	Mus muscu
36	34.2	8.6	28993	3	LMFL5808	Plasmodiu
37	34	8.6	373	3	AF221826	Leishmani
38	34	8.6	657	8	AF121142	Gracilari
39	34	8.6	7242	8	MITTARRNG	T.aestivum
40	34	8.6	103787	9	AC093659	Homo sapi
41	34	8.6	152248	2	AC027522	Homo sapi
42	34	8.6	184034	2	AC024156	Homo sapi
43	33.8	8.5	1154	4	AF127029	Bos tauru
44	33.8	8.5	52639	2	AC107472	Rattus no
45	33.8	8.5	64790	2	AC099356	Rattus no

ALIGNMENTS

RESULT	1	AX061229	Sequence	76 from Patent WO0078953.	1322 bp	DNA	linear	PAT 22-JAN-2001
LOCUS	AX061229	Sequence	76 from Patent WO0078953.					
DEFINITION	AX061229	Sequence	76 from Patent WO0078953.					
ACCESSION	AX061229	Sequence	76 from Patent WO0078953.					
VERSION	AX061229.1	Sequence	76 from Patent WO0078953.					
KEYWORDS	AX061229.1	Sequence	76 from Patent WO0078953.					
SOURCE	AX061229.1	Sequence	76 from Patent WO0078953.					
ORGANISM	AX061229.1	Sequence	76 from Patent WO0078953.					
REFERENCE	AX061229.1	Sequence	76 from Patent WO0078953.					
AUTHORS	AX061229.1	Sequence	76 from Patent WO0078953.					
TITLE	AX061229.1	Sequence	76 from Patent WO0078953.					
JOURNAL	AX061229.1	Sequence	76 from Patent WO0078953.					
FEATURES	AX061229.1	Sequence	76 from Patent WO0078953.					
source	AX061229.1	Sequence	76 from Patent WO0078953.					
BASE COUNT	AX061229.1	Sequence	76 from Patent WO0078953.					

ORIGIN

Query Match 70.7%; Score 280; DB 6; Length 1322;
 Best Local Similarity 100.0%; Pred. No. 1.2e-81;
 Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atgcagagctacagcctgacccagctgcccgcctatgcgaatgtgtgagggccctctgg 60
 |||||
 Db 186 ATGCAGAGCTACAGCTGACCCAGCTGCCGCTATGCAATGTGTGGAGGCCCTCTGG 245
 |||||

Qy 61 aggattataagaacgagggcctatgagggccatgaggggctgaacgtcacagcaaca 120
 |||||
 Db 246 AGGATTATAAGAAGGAGGGCTATGGAGGCCCATGAGGGGCTGAAGCTCACAGCAACA 305
 |||||

Qy 121 ggcgagggcctgccagcgccttatttgcctgctacgaaaagttaaaaaaacacattg 180
 |||||
 Db 306 GCGCAGGGCCTGCCAGCGCCCTTTATTTGCTGCTACGAAAAGTTAAAAAAGACATTG 365
 |||||

Qy 181 agtgatgtaattccaccctgggggccaatagccatttgcgaatgtgcccgggtgtgtg 240
 |||||
 Db 366 AGTGATGTAATCCACCTGGGGCAATAGCCATATTGCCAATGCTGCGCGCGGTGTGTG 425
 |||||

Qy 241 gcaacattacttcagtcagccatgaacccctgcggaag 280
 |||||
 Db 426 GCAACATTACTTTCATGATGACGAGCATGAACCCCTGCGGAAG 465
 |||||

RESULT 2

AF327402 1448 bp mRNA linear PRI 02-MAY-2001
 LOCUS AF327402 Homo sapiens putative mitochondrial solute carrier splice variant
 DEFINITION mRNA, complete cds, alternatively spliced, nuclear gene for
 mitochondrial product.

ACCESSION AF327402.1 GI:13926046

VERSION AF327402

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1448)

Li, F., Nikali, K., Gregan, J., Leibiger, I., Leibiger, B.,

Schweyen, R., Larsson, C. and Suomalainen, A.

Characterization of a novel human putative mitochondrial

transporter homologous to the yeast mitochondrial RNA splicing

proteins 3 and 4

FEBS Lett. 494 (1-2), 79-84 (2001)

21195335

11297739

2 (bases 1 to 1448)

Li, F., Nikali, K., Gregan, J., Leibiger, I., Leibiger, B., Schweyen, R.,

Larsson, C. and Suomalainen, A.

Direct Submission

Submitted (08-DEC-2000) Human Molecular Genetics, National Public

Health Institute, Mannerheimintie 166, Helsinki 00300, Finland

Location/Qualifiers

1. 1448

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="10"

30. 1124

/note="similar to mitochondrial RNA splicing protein 3/4;

alternatively spliced"

/codon_start=1

/product="putative mitochondrial solute carrier splice

variant"

/protein_id="AAK49519.1"

/db_xref="GI:13926047"

/translation="MELEGRGAGVAGGPAAGPGRSPGESALLDGLORVGRGAGG

EAGACRPVQRQDSGDYDALPAGATVTVHWVAGVAGLLEHCVMPPIDCVTRMQS

LQPDPAARYNLEALRIIRTEGLWRPMRGLNVTATGAGPAHALYFACYERLUKRTLS

DVTHPGNSHIANGAAGCVATLLHDAAMPAAEVVVKORMQYNSPYHRVTDVRAVQON
 EGATFYRSYTTQLTNNVFPQAIHFMTYFLQEHFNQRRYNSPSSHLSCACAGVAA
 AATTPLDVKCTLLNTLNLNSHITGHITMASAFRTVYQGVGTAYFRGVQARVY
 QIPSTAIANSVYEFKYLITKROEWRAG"

BASE COUNT 322 a 408 c 435 g 283 t
 ORIGIN

Query Match 70.7%; Score 280; DB 9; Length 1448;
 Best Local Similarity 100.0%; Pred. No. 1.3e-81;
 Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atgcagagctacagcctgacccagctgcccgcctatgcgaatgtgtgagggccctctgg 60
 |||||
 Db 327 ATGCAGAGCTACAGCTGACCCAGCTGCCGCTATGCAATGTGTGGAGGCCCTCTGG 386
 |||||

Qy 61 aggattataagaacgagggcctatgagggccatgaggggctgaacgtcacagcaaca 120
 |||||
 Db 387 AGGATTATAAGAAGGAGGGCCTATGGAGGCCCATGAGGGGCTGAACGTCACAGCAACA 446
 |||||

Qy 121 ggcgagggcctgccagcgccttatttgcctgctacgaaaagttaaaaaaacacattg 180
 |||||
 Db 447 GGCAGGCGCTGCCAGCGCCCTTTATTTGCTGCTACGAAAAGTTAAAAAAGACATTG 506
 |||||

Qy 181 agtgatgtaattccaccctgggggccaatagccatttgcgaatgtgcccgggtgtgtg 240
 |||||
 Db 507 AGTGATGTAATCCACCTGGGGCAATAGCCATATTGCCAATGCTGCGCGGTGTGTG 566
 |||||

Qy 241 gcaacattacttcagtcagccatgaacccctgcggaag 280
 |||||
 Db 567 GCAACATTACTTTCATGATGACGAGCATGAACCCCTGCGGAAG 606
 |||||

RESULT 3

HSA303077 standard; RNA; HUM; 1448 BP.

XX HSA303077

AC AJ303077;

XX AJ303077.1

SV AJ303077.1

XX 02-FEB-2001 (Rel. 56, Created)

DT 26-APR-2001 (Rel. 67, Last updated, Version 2)

XX Homo sapiens mRNA for mitochondrial RNA splicing protein 3/4 (HMRS3/4

gene), 1448 bp cDNA splice variant

DE gene), 1448 bp cDNA splice variant

XX HMRS3/4 gene; mitochondrial RNA splicing protein 3/4.

KW Homo sapiens (human)

XX Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;

OS Euthera; Primates; Catarrhini; Hominidae; Homo.

XX [1]

RP 1-1448

RA Nikali K.;

RT ;

RL Submitted (12-DEC-2000) to the EMBL/GenBank/DBJ databases.

RL Nikali K., Human Molecular Genetics, National Public Health Institute,

Mannerheimintie 166, Helsinki, FI-00300, FINLAND.

XX [2]

RA Li F., Nikali K., Gregan J., Leibiger I., Leibiger B., Schweyen R.,

Larsson C., Suomalainen A.;

RA "Characterization of a novel human putative mitochondrial transporter

homologous to the yeast mitochondrial RNA splicing proteins 3 and 4";

FEBS Lett. 494:79-84 (2001).

XX Key

Location/Qualifiers

1. 1448

/db_xref="taxon:9606"

/organism="Homo sapiens"

FT source

FT

```

CDS
30. .1124
/ gene="HMR3/4"
/ product="mitochondrial RNA splicing protein 3/4"
/ function="putative mitochondrial solute carrier"
/ protein_id="CAC27996.1"
/ translation="MELGRGAGGVAGGPAAGPRSPGESALLDGLRGVGRGAGGGE
AGACRPVRQDPDGPDEALPAGATVTHMVAGAVAGILEHCVMYPIDCVKTRMQSLQ
PDPAARYRNLEALNRIITREGLMRPMLNATGAGPAHALYFACYEKLKTLSDVI
HPGNSHTANGAAGCAVATLLHDAANPAEVVKRQMYSNPVHRVTDVCRVAVONEGAG
AFYRSYTTOLTMNVFOAIHEFTYEFLOEHEFNPNRPNPSHVLSSGACAVAAATTP
LDVCKTLNTOESLALNSHITGTMASAFRTYVQGVGTAYFRGVARVIYQIPSTA
TANSYEFKYLITKROEWRAKG"
SQ Sequence 1448 BP; 322 A; 408 C; 435 G; 283 T; 0 other;

Query Match 70.7%; Score 280; DB 17; Length 1448;
Best Local Similarity 100.0%; Pred. No. 1.3e-81;
Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgcagagtctacagctgaccacagctgccgctatcgcaatgtgttgaggccctctgg 60
|||||
DB 327 ATGCAGAGTCTACAGGCTGACCCAGCTGCCGCTATCGCAATGTGTGGAGGCCCTCTGG 386

QY 61 agattataagaacgagggccctatggagggccatgaggggctgaacgtcacagcaaca 120
|||||
DB 387 AGGATTATAGAAGCGAGGGCCCTATGGAGGCCCATGAGGGGCTGAACGTACAGCAACA 446

QY 121 ggcgagggcctgccacccctttatttgcctgctacgaaagttaaaaagacattg 180
|||||
DB 447 GCGCAGGGCCCTGCCACGCCCTTTATTGTGCTGCTACGAAAGTTAAAAAGACATTG 506

QY 181 agtgatgtaatccacccctggggcaatagccattgccaatggtgcggcggtgtgtg 240
|||||
DB 507 AGTGATGTAATCCACCCTGGGGCAATACCCATATTGCCAATGGTGGCGCGGTGTGTG 566

QY 241 gaacattacttcattgacgacctgaacctgcggaag 280
|||||
DB 567 GCAACATTACTTCATGATGATGACCCATGAACCCCTGCGGAAG 606

RESULT 4
AX071440
LOCUS AX071440 401 bp DNA linear PAT 25-JAN-2001
DEFINITION Sequence 1912 from Patent WO0102568.
ACCESSION AX071440
VERSION AX071440.1 GI:12581791
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 401)
AUTHORS Williams,L.T., Escobedo,J., Innis,M.A., Garcia,P.D., Klinger,J.,
Kassam,A., Reinhard,C., Randazzo,F., Kennedy,G.C., Pot,D.,
Lanson,G., Drmanac,R., Crkenjakov,R., Drmanac,S., Dickson,M.,
Labat,I., Leshkowitz,D., Kita,D., Garcia,V. and Strache-Crain,B.
TITLE Human genes and gene expression products
JOURNAL Patent: WO 0102568-A 1912 11-JAN-2001;
CHIRON CORPORATION (US) ; HYSEQ, INC. (US)
FEATURES
Location/Qualifiers
source 1..401
/organism="Homo sapiens"
/db_xref="taxon:9606"
misc_feature 1..401
/ note="n = A,T,C or G"
BASE COUNT 100 a 104 c 120 g 76 t 1 others
ORIGIN

Query Match 68.9%; Score 273; DB 6; Length 401;
Best Local Similarity 100.0%; Pred. No. 2.1e-79;
Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```

```

QY 8 gtctacagctgacccagctgccgctatcgcaatgtgttgaggccctctgaggatta 67
|||||
DB 9 GTCTACAGGCTGACCCAGCTGCCGCTATCGCAATGTGTGGAGGCCCTCTCGAGGATTA 68

QY 68 taagaacgagggccctatggagggccatgaggggctgaacctcacagcaacagcgcgag 127
|||||
DB 69 TAAGAACGAGGGCCCTATGAGGCCCATGAGGGGCTGAACGTTCACAGCAACAGCGCGAG 128

QY 128 ggcctgccacccctttatttgcctgctacgaaagttaaaaagacattgagtgtg 187
|||||
DB 129 GGCCTGCCACGCCCTTTATTGTGCTGCTACGAAAGTTAAAAAGACATTGAGTGATG 188

QY 188 taatccacctgggggcaatagccattgccaatggtgcggcggtgtgtggaacat 247
|||||
DB 189 TAATCCACCTGGGGCAATAGCCATATTGCCAATGGTGGCGCGGTGTGTGGCAACAT 248

QY 248 tacttcattgacgacctgaacctgcggaag 280
|||||
DB 249 TACTTCATGATGACGCATGAACCCCTGCGGAAG 281

RESULT 5
AF327403
LOCUS AF327403 1889 bp mRNA linear PRI 02-MAY-2001
DEFINITION Homo sapiens putative mitochondrial solute carrier splice variant
mRNA, complete cds, alternatively spliced, nuclear gene for
mitochondrial product.
ACCESSION AF327403
VERSION AF327403.1 GI:13926049
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1889)
AUTHORS Li,F.Y., Nikali,K., Grogan,J., Leibiger,I., Leibiger,B.,
Schweyen,R., Larsson,C. and Suomalainen,A.
TITLE Characterization of a novel human putative mitochondrial
transporter homologous to the yeast mitochondrial RNA splicing
proteins 3 and 4
JOURNAL FEBS Lett. 494 (1-2), 79-84 (2001)
MEDLINE 21195335
PUBMED 11297739
REFERENCE 2 (bases 1 to 1889)
AUTHORS Li,F., Nikali,K., Grogan,J., Leibiger,I., Leibiger,B., Schweyen,R.,
Larsson,C. and Suomalainen,A.
TITLE Direct Submission
JOURNAL Submitted (08-DEC-2000) Human Molecular Genetics, National Public
Health Institute, Mannerheimintie 166, Helsinki 00300, Finland
FEATURES
Location/Qualifiers
source 1..1889
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10"
/map="10q24"
1032..1565
/ note="similar to mitochondrial RNA splicing protein 3/4;
alternatively spliced"
/codon_start=1
/product="putative mitochondrial solute carrier splice
variant"
/protein_id="AAK49520.1"
/db_xref="GI:13926050"
/translation="MNPAAVYKQRMQYNSPHYRVDVCVRVWONEGAGAFYRSYTTQ
LTMNVPFOAIHEFTYEFLOEHEFNPNRPNPSHVLSSGACAVAAATTPIDVCKTL
NTQESLALNSHITGTMASAFRTYVQGVGTAYFRGVARVIYQIPSTAIANSVTE
FEKYLITKROEWRAKG"
BASE COUNT 404 a 508 c 544 g 433 t
ORIGIN

Query Match 56.8%; Score 225; DB 9; Length 1889;


```

```

Best Local Similarity 100.0%; Pred. No. 2.5e-63;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atgcagagctacagcctgaccagctgcccgcctatgcgaatgtgttgaggccctctgg 60
Db 306 ATGCAGAGCTACAGCCTGACCCAGCTGCCCGCTATPCGAATGTGTGGAGGCCCTCTGG 365
Qy 61 aggattataagaacgagggcctatggagcccatgaggggctgaagcgtcacagcaaca 120
Db 366 AGGATTATAAGAACGAGGGCCTATGGAGGCCATGAGGGGCTGAACGTCACACCAACA 425
Qy 121 ggcgagggcctgcacagccctttatttgcctgctacgaaagttaaaaaagacattg 180
Db 426 GCGCAGGGCCTGCCACGCCCTTTATTTGCTGCTACGAAAGTTAAAAAGACATTG 485
Qy 181 agtgaatgaatccaccctgggggcaatgacgaattgccaatggt 225
Db 486 AGTGATGTAATCCACCTGGGGGCAATAGCCATATTGCCAATGGT 530

RESULT 6
HSA303078
ID HSA303078 standard; RNA; HUM; 1889 BP.
AC AJ303078;
SV AJ303078.1
XX
DT 02-FEB-2001 (Rel. 66, Created)
DT 26-APR-2001 (Rel. 67, Last updated, Version 2)
XX
DE Homo sapiens mRNA for mitochondrial RNA splicing protein 3/4 (HMR3/4
DE gene), 1889 bp cDNA splice variant
XX
KW HMR3/4 gene; mitochondrial RNA splicing, protein 3/4.
XX
OS Homo sapiens (human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
XX
[1]
RN 1-1889
RA Nikali K.;
RT ;
RL Submitted (12-DEC-2000) to the EMBL/GenBank/DBJ databases.
RL Nikali K., Human Molecular Genetics, National Public Health Institute,
RL Mannerheimintie 166, Helsinki, FI-00300, FINLAND.
XX
[3]
RA Li F., Nikali K., Gregan J., Leibiger I., Leibiger B., Schweyen R.,
RA Larsson C., Suomalainen A.;
RT "Characterization of a novel human putative mitochondrial transporter
RT homologous to the yeast mitochondrial RNA splicing proteins 3 and 4";
RT FEBS Lett. 494:79-84(2001).
XX
FH Key Location/Qualifiers
FT 1..1889
FT /db_xref="taxon:9606"
FT /organism="Homo sapiens"
FT CDS
FT 1032..1565
FT /gene="HMR3/4"
FT /product="mitochondrial RNA splicing protein 3/4"
FT /function="putative mitochondrial solute carrier"
FT /protein_id="CAC27997.1"
FT /translation="MNPAAVVKRMQMYNSPHRVTDICRAVWQNEGAGAFYRSYTTQL
FT TMNVPFQAIIHMTYFTLQEHENPQRRYNPSHVLGACAGAAAAATPLDVKTLTNT
FT QESLALNSHITGHTGMASAPRTYVQGGVTAYFRGVQARVIYQIPSTAIWSVYEFK
FT YLITKQEWNRAG"
XX
SQ Sequence 1889 BP; 404 A; 508 C; 544 G; 433 T; 0 other;

```

```

Query Match 56.8%; Score 225; DB 17; Length 1889;
Best Local Similarity 100.0%; Pred. No. 2.5e-63;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atgcagagctacagcctgaccagctgcccgcctatgcgaatgtgttgaggccctctgg 60
Db 306 ATGCAGAGCTACAGCCTGACCCAGCTGCCCGCTATPCGAATGTGTGGAGGCCCTCTGG 365
Qy 61 aggattataagaacgagggcctatggagcccatgaggggctgaagcgtcacagcaaca 120
Db 366 AGGATTATAAGAACGAGGGCCTATGGAGGCCATGAGGGGCTGAACGTCACACCAACA 425
Qy 121 ggcgagggcctgcacagccctttatttgcctgctacgaaagttaaaaaagacattg 180
Db 426 GCGCAGGGCCTGCCACGCCCTTTATTTGCTGCTACGAAAGTTAAAAAGACATTG 485
Qy 181 agtgaatgaatccaccctgggggcaatgacgaattgccaatggt 225
Db 486 AGTGATGTAATCCACCTGGGGGCAATAGCCATATTGCCAATGGT 530

RESULT 7
AL353719/c
LOCUS
DEFINITION Human DNA sequence from clone RP11-85A1 on chromosome 10, complete
sequence.
ACCESSION AL353719 AC007643
VERSION AL353719.10 GI:15787725
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 123160)
Direct Submission
Submitted (25-SEP-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Sep 26, 2001 this sequence version replaced gi:14280413.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em; EMBL; SW;
SWISSPROT; Tr; TREMBL; Wp; WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
This sequence
was generated from part of bacterial clone contigs of human
chromosome 10, constructed by the Sanger Centre Chromosome 10
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr10
RP11-85A1 is from the library RPI-11.1 constructed by the group of
Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBAC3.6
IMPORTANT: This sequence is not the entire insert of clone
RP11-85A1. It may be shorter because we sequence overlapping
sections only once, except for a short overlap.
The true left end of clone RP11-85A1 is at 1 in this sequence. The
true left end of clone RP11-483F11 is at 123061 in this sequence.
The true right end of clone RP11-129J12 is at 51589 in this
sequence.

```

```

FEATURES
  source      Location/Qualifiers
1. .123160
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /chromosome="10"
   /clone="RP11-85A1"
   /clone_lib="RPC1-11.1"
   /complement(7066..7118)
   /note="Single clone region. Sequence from reads from a
   short insert library derived from a single pUC clone.
   Restriction digest data confirm the assembly."
52162
   52162
   52167
   52167
   33221 a 28084 c 27196 g 34659 t
BASE COUNT
ORIGIN

Query Match      56.8%; Score 225; DB 9; Length 123160;
Best Local Similarity 100.0%; Pred. No. 5.le-63;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atcgagctctacagcctgaccagctccgcctatcgaatgtgttgagccctctgg 60
|||||
Db 85540 ATGCAGAGTCTACACCTGACCCAGCTCCCGCTATCGCAATGTTGTGAGCGCCCTCTGG 85481
|||||

Qy 61 aggtattataagaagcgggctctgagggccatgagggccatgaggggctgaacgtcacagcaaca 120
|||||
Db 85480 AGGATTATAAGAACGAGGCGCTATGAGGCGCCATGAGGGGCTGACGTCACACCAACA 85421
|||||

Qy 121 ggcgcagggcctgcccacgcccctttatttgcctgctacgaaaagttaaaaagacattg 180
|||||
Db 85420 GCGCAGAGCGCTGCGCACGCGCCCTTTATTTCCTGCTACGAAAAGTTAAAAAGACATTG 85361
|||||

Qy 161 agtgaatgaatccacctgggggcaatagcccatattgccaatggt 225
|||||
Db 85360 ACTGATGTAATCCACCTGGGGGCAATAGCCCATATFTGCCAATGGT 85316
|||||

RESULT 8
AC096351/c
LOCUS
DEFINITION Rattus norvegicus chromosome R1 clone CH230-24M6, WORKING DRAFT
SEQUENCE, 33 unordered pieces.
ACCESSION AC096351
VERSION AC096351.2 GI:17944054
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 221062)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.B., Amaratunge,H.C., Are,J.R., Banks,T., Barbaria,J.,
Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowle,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle-McRroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
Hollins,B., Homsí,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,

```

Loulseged, H., Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R.,
 Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A.,
 Martinez, E., Massey, E., Mawhney, E., McLeod, M.P., Meador, M.,
 Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K.,
 Morgan, A., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N.,
 Nguyen, A., Nguyen, N., Nickerson, N., Nickerson, E., Nwokenkwo, S.,
 Oguh, M., Okwuonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B.,
 Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L.,
 Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojubokan, I., Rolfe, M.,
 Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoohtari, N.,
 Sisson, I., Sodergren, E., Sonaik, T., Sparks, A., Stanley, H.,
 Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K.,
 Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N.,
 Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R.,
 Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C.,
 Watlington, S., Williams, G., Williamson, A., Wleczek, R., Wooden, S.,
 Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
 Weinstein, G. and Gibbs, R.
 Direct Submission
 Unpublished
 2 (bases 1 to 221062)
 Worley, K.C.
 Direct Submission
 Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Dec 20, 2001 this sequence version replaced gi:15627972.
 ----- Genome Center of Medicine
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GE2D
 Center clone name: CH230-24M6
 ----- Summary Statistics
 Assembly program: Phrap; version 0.990329First call to
 findPhrapList
 Consensus quality: 200164 bases at least Q40
 Consensus quality: 205303 bases at least Q30
 Consensus quality: 209366 bases at least Q20
 Estimated insert size: 209880; sum-of-contigs estimation
 Quality coverage: 0x in Q20 bases; agarose-fp estimation
 Quality coverage: 3.9x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 33 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 32851: contig of 32851 bp in length
 * 32852 32951: gap of unknown length
 * 32952 48984: contig of 16033 bp in length
 * 48985 49084: gap of unknown length
 * 49085 63815: contig of 14731 bp in length
 * 63816 63915: gap of unknown length
 * 63916 76005: contig of 12090 bp in length
 * 76006 76105: gap of unknown length
 * 76106 88841: contig of 12736 bp in length
 * 88842 88941: gap of unknown length
 * 88942 100591: contig of 11650 bp in length
 * 100592 100691: gap of unknown length
 * 100692 111702: contig of 11011 bp in length
 * 111703 111802: gap of unknown length
 * 111803 122664: contig of 10862 bp in length
 * 122665 122764: gap of unknown length
 * 122765 129711: contig of 6947 bp in length
 * 129712 129811: gap of unknown length


```

RESULT 10
AF223466
LOCUS      Homo sapiens HT015 protein (HT015) mRNA, complete cds.
DEFINITION
ACCESSION AF223466
VERSION   AF223466.1 GI:7578782
KEYWORDS
SOURCE    human.
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS   Song, H., Gao, G., Peng, Y., Ren, S., Chen, Z. and Han, Z.
TITLE     A novel gene expressed in human hypothalamus
JOURNAL   Unpublished
REFERENCE
AUTHORS   Song, H., Gao, G., Peng, Y., Ren, S., Chen, Z. and Han, Z.
TITLE     Direct Submission
JOURNAL   Submitted (12-JAN-2000) Chinese National Human Genome Center at
           Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong,
           Shanghai 201203, China
FEATURES
Source    1..1429
           /organism="Homo sapiens"
           /db_xref="taxon:9606"
           /tissue_type="hypothalamus"
gene      1..1429
           /gene="HT015"
CDS       45..1088
           /gene="HT015"
           /codon_start=1
           /product="HT015 protein"
           /protein_id="AAF64141.1"
           /db_xref="GI:7578783"
           /translation="MELRSSGVSQAVARMGDSRGGGKGDKATGSGEDYENLPTSS
           VSHMTAGAMGILIEVMPVDSVKTRMQSLSPDKAQYTSIYGALKKINRTGFWR
           PLRVNYIMGLQAPAHMYFAICENMKRTLNDVPHQGNSHLANGSAGSMATLLHDVA
           MNPAEYVVKQMSNRSAISCIRTWRTGELGAFYRSYTTOLTWNIPQSIHFIT
           YEFQEOVNPHTVPQSHIISGGLAGALAAATPLDVCCKTLTNTQENVALSIANTIS
           GRLSGMANFRTVYQNLPLPATSKASRVSTRCPFPFGLSLSSSSTSPSPASWKI
           ELHTKGRDHRIFS"
BASE COUNT 326 a 426 c 402 g 275 t
ORIGIN

Query Match      32.3%; Score 128; DB 9; Length 1429;
Best Local Similarity 66.1%; Pred. No. 4.1e-31;
Matches 185; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

Qy 1 atcgagagctacagcctgacccagctgcccgcctatcgcaatgtgttgaggccctctgg 60
Db 261 ATGCAGAGTTTGAGTCCAGATCCCAAGCCAGTACACAAGTATCTACGGAGCCCTCAAG 320

Qy 61 aggattataagacgagggcctatggaggccatgaggggctgaacgtcacagcaaca 120
Db 321 AAAATCATGCGGACGGAAGGCTTCTGGAGGCCCTTGGAGGCGCTCAACGTCATGATCATG 380

Qy 121 ggcgcaggcctgcccgcgccttatttgcctgctacgaaagttaaagaaagcattg 180
Db 381 GGTGCGAGCGCGGCCCATGACCATGATATTTTTCCTGCTATGAAACATGAAAGGACTTTA 440

Qy 181 agtgatgtaatccaccctgggggcaatagccatattgccaatggcgccgggtgtgtg 240
Db 441 AATGACGTTTTCCACCACCAAGAAACAGCCACCCTAGCACGGGATAGCTGGGAGTATG 500

Qy 241 gcaacattacttcattgatgcagccatgaacccctcgaggag 280
Db 501 GCCACCCCTGCTCCACGATCGCGTATGATGAATCCAGCAGAG 540

LOCUS      AC099368      163205 bp      DNA      linear      HTG 20-DEC-2001
DEFINITION Rattus norvegicus clone CH230-84G11, *** SEQUENCING IN PROGRESS
***, 71 unordered pieces.
ACCESSION AC099368
VERSION   AC099368.2 GI:17973175
KEYWORDS HTG; HTGS_PHASE1.
SOURCE    Norway rat.
ORGANISM  Rattus norvegicus
           Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
           Rattus.
REFERENCE
AUTHORS   1 (bases 1 to 163205)
           Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C.,
           Albrooks, S.L., Amaratunge, H.C., Are, J.R., Banks, T., Barbaria, J.,
           Benton, J., Bimage, K., Blankenburg, K., Bonnin, D., Bouck, J.,
           Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C.,
           Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F.,
           Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R.,
           Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C.,
           Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C.,
           Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O.,
           Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H.,
           Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C.,
           Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J.,
           Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T.,
           Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S.,
           Hamilton, K., Harris, K., Harris, K., Hart, M., Havlak, P., Hawes, A.,
           Hernandez, J., Hernandez, O., Hodgson, A., Hoques, M., Holloway, C.,
           Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J.,
           Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S.,
           Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J.,
           Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C.,
           Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W.,
           Loulseg, H., Lozado, R. J., Lu, X., Lucier, A., Lucier, R., Luna, R.,
           Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A.,
           Martinez, E., Massey, E., Mawhiney, E., McLeod, M.P., Meador, M.,
           Mel, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K.,
           Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N.,
           Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S.,
           Ogih, M., Okwuonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B.,
           Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L.,
           Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojubokan, I., Rolfe, M.,
           Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoohtari, N.,
           Slison, I., Sodergren, E., Sonalike, T., Sparks, A., Stanley, H.,
           Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K.,
           Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N.,
           Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R.,
           Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C.,
           Watlington, S., Williams, G., Williamson, A., Wleczyk, R., Wooden, S.,
           Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
           Weinstein, G. and Gibbs, R.
REFERENCE
AUTHORS   2 (bases 1 to 163205)
           Direct Submission
           Worley, K.C.
TITLE     Submitted (10-NOV-2001) Human Genome Sequencing Center, Department
           of Molecular and Human Genetics, Baylor College of Medicine, One
           Baylor Plaza, Houston, TX 77030, USA
COMMENT   On Dec 20, 2001 this sequence version replaced gi:16901901.
           ----- Genome Center
           Center: Baylor College of Medicine
           Center code: BCM
           Web site: http://www.hgsc.bcm.tmc.edu/
           Contact: hgsc-help@bcm.tmc.edu
           ----- Project Information
           Center project name: GH1J
           Center clone name: CH230-84G11
           ----- Summary Statistics
           Assembly program: Phrap; version 0.990329First call to
           findPhrapList
           Consensus quality: 127058 bases at least Q40
           Consensus quality: 138264 bases at least Q30

```

Consensus quality: 146548 bases at least Q20
Estimated insert size: 120308; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 1.9x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 71 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 8509: contig of 8509 bp in length
* 8510: gap of unknown length
* 8610: contig of 4931 bp in length
* 13541: 13640: gap of unknown length
* 13641: 19516: contig of 5876 bp in length
* 19517: 19616: gap of unknown length
* 19617: 24900: contig of 5374 bp in length
* 24991: 25090: gap of unknown length
* 25091: 29401: contig of 4311 bp in length
* 29402: 29501: gap of unknown length
* 29502: 33153: contig of 3552 bp in length
* 33054: 33154: gap of unknown length
* 33154: 38439: contig of 5286 bp in length
* 38440: 38539: gap of unknown length
* 38540: 41183: contig of 2644 bp in length
* 41184: 41283: gap of unknown length
* 44936: 45036: contig of 3653 bp in length
* 45037: 47637: contig of 2601 bp in length
* 47638: 47737: gap of unknown length
* 47738: 50166: contig of 2429 bp in length
* 50167: 50266: gap of unknown length
* 50267: 54170: contig of 3904 bp in length
* 54171: 54270: gap of unknown length
* 54271: 57672: contig of 3402 bp in length
* 57672: 57772: gap of unknown length
* 57773: 61162: contig of 3390 bp in length
* 61163: 61263: gap of unknown length
* 61263: 63890: contig of 2628 bp in length
* 63891: 63990: gap of unknown length
* 63991: 66900: contig of 2910 bp in length
* 66901: 67000: gap of unknown length
* 67001: 69717: contig of 2717 bp in length
* 69718: 69817: gap of unknown length
* 69818: 72921: contig of 3104 bp in length
* 72922: 73021: gap of unknown length
* 73022: 75145: contig of 2125 bp in length
* 75147: 75246: gap of unknown length
* 75247: 79038: contig of 3792 bp in length
* 79039: 79138: gap of unknown length
* 79139: 80340: contig of 1202 bp in length
* 80341: 80440: gap of unknown length
* 80441: 82366: contig of 1926 bp in length
* 82367: 82466: gap of unknown length
* 82467: 84184: contig of 1718 bp in length
* 84185: 84284: gap of unknown length
* 84285: 86001: contig of 1717 bp in length
* 86002: 86101: gap of unknown length
* 86102: 87695: contig of 1594 bp in length
* 87696: 87795: gap of unknown length
* 87796: 89287: contig of 1492 bp in length
* 89288: 89387: gap of unknown length
* 89388: 91950: contig of 2563 bp in length
* 91951: 92050: gap of unknown length
* 92051: 94575: contig of 2525 bp in length
* 94576: 94673: gap of unknown length
* 94674: 96581: contig of 1906 bp in length
* 96582: 96681: gap of unknown length

96682 98235: contig of 1554 bp in length
* 98236: gap of unknown length
* 98336 98814: contig of 1479 bp in length
* 99815 99914: gap of unknown length
* 100964: contig of 1050 bp in length
* 100965 101064: gap of unknown length
* 101065 103074: contig of 2010 bp in length
* 103075 103174: gap of unknown length
* 103175 104311: contig of 1137 bp in length
* 104312 104411: gap of unknown length
* 104412 106834: contig of 2423 bp in length
* 106835 106934: gap of unknown length
* 106935 108922: contig of 1988 bp in length
* 108923 109022: gap of unknown length
* 109023 110868: contig of 1846 bp in length
* 110869 110968: gap of unknown length
* 110969 113056: contig of 2088 bp in length
* 113057 113156: gap of unknown length
* 113157 114296: contig of 1140 bp in length
* 114297 114396: gap of unknown length
* 114397 116312: contig of 1916 bp in length
* 116313 116412: gap of unknown length
* 116413 117492: contig of 1080 bp in length
* 117493 117592: gap of unknown length
* 117593 119552: contig of 1960 bp in length
* 119553 119652: gap of unknown length
* 119653 120658: contig of 1006 bp in length
* 120659 120758: gap of unknown length
* 120759 121895: contig of 1137 bp in length
* 121896 121995: gap of unknown length
* 121996 123272: contig of 1277 bp in length
* 123273 123372: gap of unknown length
* 123373 125310: contig of 1938 bp in length
* 125311 125410: gap of unknown length
* 125411 126891: contig of 1481 bp in length
* 126892 126991: gap of unknown length
* 126992 128690: contig of 1699 bp in length
* 128691 128790: gap of unknown length
* 128791 130827: contig of 2037 bp in length
* 130828 130927: gap of unknown length
* 130928 132501: contig of 1574 bp in length
* 132502 132601: gap of unknown length
* 132602 134071: contig of 1470 bp in length
* 134072 134171: gap of unknown length
* 134172 135838: contig of 1667 bp in length
* 135839 135938: gap of unknown length
* 135939 137589: contig of 1651 bp in length
* 137590 137689: gap of unknown length
* 137690 139043: contig of 1354 bp in length
* 139044 139143: gap of unknown length
* 139144 140777: contig of 1634 bp in length
* 140778 140877: gap of unknown length
* 140878 142526: contig of 1649 bp in length
* 142527 142626: gap of unknown length

Query Match 31.6%; Score 125.2; DB 2; Length 163205;
Best Local Similarity 87.4%; Pred. No. 7.9e-30;
Matches 194; Conservative 0; Mismatches 23; Indels 5; Gaps 5;

QY 7 agtctaacgctgaccag -ctgccgctatcgcaatgtgtgga-ggcccctctggagga 64
Db 13720 ATTTTACAGCTGATCCAGCGCGCCCTTATCGAAAAAGTTTGAGGGCTTTTGGAGAA 13779
QY 65 ttataagaacggaggcctatggaggcccatga-gggggcgtgaacgtcacagcaacagc 123
Db 13780 TTATAAGAACAGAGCGCCCTGTGGAGGCCCATGAGGGGGTTGACGTCACCTGCCTGGC 13839
QY 124 gcaggcctgcccacgccccttatttgcctgcacgaaaaagttaaaaaagacattagat 183
Db 13840 GCGGGGCTGCCACCGCCCTCTATTG-CTGCTACGAAAAAG-TAAAAAGACATTGAGT 13897
QY 184 gatgtaatccacccctgggggcaatagccatatgtccaatggt 225
|||||

Db 13898 GATGTAATCCACCCGGGGCAATAGCCATATTGGCAATGGT 13939

RESULT 12

AF361699 5869 bp mRNA linear ROD 07-NOV-2001

LOCUS

DEFINITION

AF361699 Mus musculus mitochondrial carrier-like protein mRNA, complete cds;

AF361699 nuclear gene for mitochondrial product.

AF361699.1 GI:16755527

house mouse.

SOURCE

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 5869)

Li,Q., Eckenrode,S., Wang,C., Ruan,Q., Shi,J., McIndoe,R.A. and She,J.

A novel mouse mitochondrial carrier protein gene is up-regulated from young to adult NOD mice

Unpublished

2 (bases 1 to 5869)

Li,Q., Eckenrode,S., Wang,C., Ruan,Q., Shi,J., McIndoe,R.A. and She,J.

Direct Submission

Submitted (15-MAR-2001) Department of Pathology, Immunology and Laboratory Sciences, University of Florida, 1600 SW Archer Road, Room D6-15, Gainesville, FL 32610, USA

Location/Qualifiers

1..5869

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/chromosome="14"

/map="between D14Mit124 and D14Mit236"

116..664

/note="MCLP"

/codon_start=1

/evidence=not_experimental

/product="mitochondrial carrier-like protein"

/protein_id="AAI27990.1"

/db_xref="GI:16755528"

/translation="MELRRGVGNCAAGRRMDGCRGGCGSKDAGEDYENLPTNS VSTRGLNMMGAGPAHAMFYACYENMKRTLNDVFSHQGNHLANGILKAFVWSWEALL SGASSPGPSNLHPROTENSRYT"

BASE COUNT 1387 a 1427 c 1390 g 1665 t

ORIGIN

Query Match 27.7%; Score 109.8; DB 10; Length 5869;

Best Local Similarity 68.0%; Pred. No. 5.9e-25;

Matches 153; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 1 atcagagctacagcctgaccagctgccgcctatcgcaatgtgttgaggccctctgg 60

||||| | | | | | | | | | | | | | | | | | | | | |

Db 332 ATGACAGATTGAATCCAGATCCCAAGCCCGGTATACAAGCATCTATGGCGCCCTCAAG 391

||||| | | | | | | | | | | | | | | | | | | | | |

QY 61 aggatataagaagcggcctatgagcccatgagggcctgaacgtcacagcaaca 120

||||| | | | | | | | | | | | | | | | | | | | | |

Db 392 AGATCATGCACACTGAAGGCTTCTGGAGGCCCTCGCGGGCTCGAAGCTGATGATG 451

||||| | | | | | | | | | | | | | | | | | | | | |

QY 121 ggcgcaggcctgccacgccctttatttgcctgctacgaaagttaaaaaagacattg 180

||||| | | | | | | | | | | | | | | | | | | | | |

Db 452 GGTGCAGGCGCCGCACGCCATGATTGCTGCTATGAACAACATGAAAGGACTTTA 511

||||| | | | | | | | | | | | | | | | | | | | | |

QY 181 agtgatgtaatccacctctggggcaatgacccatattgcaatgg 225

||||| | | | | | | | | | | | | | | | | | | | | |

Db 512 AATGACGCTTTTACGCCCAACGAAGAAACAGCCCATAGCTAATGTT 556

||||| | | | | | | | | | | | | | | | | | | | | |

RESULT 13

AC108878 191702 bp DNA linear HTG 06-FEB-2000

Mus musculus clone RP23-256H24 strain C57BL6/J, *** SEQUENCING IN PROGRESS ***, 17 unordered pieces.

AC108878

AC108878.2 GI:18542954

HTG; HTGS_PHASE1.

house mouse.

SOURCE

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 191702)

Grills,G., Li,L., Montgomery,K.T., Brown,W.A., Chiu,D., Decker,J., Fusina,M., Haider,A., Keller,A., Perera,A., Shim,C., Thomas,E., Zenchek,W., Xi,C., Juels,P. and Kucherlapati,R.

High Throughput Mouse Sequencing

Unpublished

2 (bases 1 to 191702)

Grills,G., Li,L., Montgomery,K.T., Brown,W.A., Chiu,D., Decker,J., Fusina,M., Haider,A., Keller,A., Perera,A., Shim,C., Thomas,E., Zenchek,W., Xi,C., Juels,P. and Kucherlapati,R.

Direct Submission

Submitted (01-FEB-2002) Harvard Partners Center for Genetics and Genomics, Harvard Medical School, 65 Landsdowne St, Cambridge, MA 02139, USA

On Feb 6, 2002 this sequence version replaced gi:18464039.

-----Genome Center

Center: Harvard Partners Genome Center

Center Code: HPGC

Web site:

http://wchanning.bwh.harvard.edu:9088/hpcgg/jsp/hpcgg/Sequence/mouse.html

Contact: gnktm@capcod.bwh.harvard.edu

-----Summary Statistics

Center project name: AKY

Sequencing vector: pUC18: L08752

Chemistry: Dye-terminator Big Dye; 100%

*Consensus quality: 187211 at least Q20

*Consensus quality: 185727 at least Q30

*Consensus quality: 183181 at least Q40

Estimated insert size: agarose-PP - N/A

**Estimated insert size: 191382 - sum-of-contigs

Quality coverage: agarose-PP - N/A

Quality coverage: 5.6 x in Q20 bases; sum-of-contigs estimation

* NOTE: This is a 'working draft' sequence. It currently

* consists of 17 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 1 48676: contig of 48676 bp in length

* 48677 48896: gap of unknown length

* 48697 89065: contig of 40369 bp in length

* 89066 89085: gap of unknown length

* 89086 104994: contig of 15909 bp in length

* 104995 105014: gap of unknown length

* 119611 119612: contig of 14598 bp in length

* 119613 119632: gap of unknown length

* 119633 135969: contig of 16337 bp in length

* 135970 135989: gap of unknown length

* 135990 151803: contig of 15814 bp in length

* 151804 151823: gap of unknown length

* 151824 160052: contig of 8229 bp in length

* 160053 160072: gap of unknown length

* 160073 169021: contig of 8949 bp in length

* 169022 169041: gap of unknown length

* 169042 175745: contig of 6704 bp in length

* 175746 175765: gap of unknown length

* 175766 180321: contig of 4556 bp in length

* 180322 180341: gap of unknown length

* 180342 182588: contig of 2247 bp in length

182589 182608: gap of unknown length
* 182609 183975: contig of 1367 bp in length
* 183976 183995: gap of unknown length
* 183996 186393: contig of 2398 bp in length
* 186394 186413: gap of unknown length
* 186414 188416: contig of 2003 bp in length
* 188417 188436: gap of unknown length
* 188437 189584: contig of 1148 bp in length
* 189585 189604: gap of unknown length
* 189605 190648: contig of 1044 bp in length
* 190649 190668: gap of unknown length
* 190669 191702: contig of 1034 bp in length.
FEATURES
source
1..191702
/organism="Mus musculus"
/strain="C57BL6/J"
/db_xref="taxon:10090"
/clone="RP23-256H24"
/sex="male"
BASE COUNT 53763 a 43806 c 42665 g 51108 t 360 others
ORIGIN
Query Match 27.7%; Score 109.8; DB 2; Length 191702;
Best Local Similarity 68.0%; Pred. No. 1.1e-24;
Matches 153; Conservative 0; Mismatches 72; Indels 0; Gaps 0;
Qy 1 atcagagctctacgctgaccgctgacccgctgcccgcctatcgcaatgtgttgaggccctctgg 60
Db 106283 ATGCAGAGTTTGAATCCAGATCCCAAGCCGATATCAAGCATCTATGGCGCCTCAAG 106342
Qy 61 aggtattataagacgaggccctatgagccctatgagccctatgagggcctgaaacgctcacagaaca 120
Db 106343 AGGATCATGCACATGAAGCTTCTGAGGCCCCCTCGGGCCCTGAACGATGATGATG 106402
Qy 121 ggcgagggcctgcccgcctctatttgcctgctacgaaagttaaagacacattg 180
Db 106403 GGTGACGGCCCGGCACGCCATGATTTTGGCTGCTATGAACATGAAAGGACTTTA 106462
Qy 181 agtattatcaatccacctgggggcaatgacccatattgccaatggt 225
Db 106463 AATGACGTTTTACGCCCAACCAAGGAACAGCCATCTAGCTAATGTT 106507
RESULT 14
AF216674
LOCUS
DEFINITION Homo sapiens chromosome 8 clone RPI-158P13 map 8p, WORKING DRAFT
SEQUENCE, 16 unordered pieces.
ACCESSION AF216674
VERSION AF216674.3 GI:14280182
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 108765)
AUTHORS Schilhabel,M.B., Baumgart,C., Blechschmidt,K., Dette,M., Jahn,N.,
Lehmann,R., Menzel,U., Polley,A., Reichwald,K., Schudy,A.,
Siddiqui,R., Taudien,S., Wen,G., Schlegelberger,B., Siebert,R.,
Rosenthal,A. and Platzzer,M.
Chromosome 8 genomic sequence
Unpublished
TITLE 2 (bases 1 to 108765)
REFERENCE Polley,A., Wen,G., Baumgart,C., Dette,M., Jahn,N., Schilhabel,M.,
Menzel,U. and Rosenthal,A.
AUTHORS
TITLE Direct Submission
JOURNAL Submitted (15-DEC-1999) Genome Analysis, Institute of Molecular
Biotechnology, Reutenbergstrasse 11, Jena 07745, Germany
COMMENT On Jun 2, 2001 this sequence version replaced gi:8151796.
----- Genome Center
Center: Institute of Molecular Biotechnology
Center code: IMB

Web site: http://genome.imb-jena.de/
Contact: gscj-submit@genome.imb-jena.de
----- Project Information
Center project name: H241
Center clone name: RPI-158P13
----- Summary Statistics
Sequencing vector: M13; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 87772 bases at least Q40
Consensus quality: 95159 bases at least Q30
Consensus quality: 100612 bases at least Q20
Quality coverage: 3,95 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 16 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1101: contig of 1101 bp in length
* 1201: gap of unknown length
* 1202 21139: contig of 19938 bp in length
* 21140 21239: gap of unknown length
* 21240 36201: contig of 14962 bp in length
* 36202 36301: gap of unknown length
* 36302 47101: contig of 10800 bp in length
* 47102 47201: gap of unknown length
* 47202 56806: contig of 9605 bp in length
* 56807 56907: gap of unknown length
* 56908 65083: contig of 8177 bp in length
* 65084 65183: gap of unknown length
* 65184 73521: contig of 8338 bp in length
* 73522 73622: gap of unknown length
* 73623 81423: gap of unknown length
* 81424 86503: contig of 5080 bp in length
* 86504 86603: gap of unknown length
* 86604 89473: contig of 2870 bp in length
* 89474 89573: gap of unknown length
* 89574 92369: contig of 2796 bp in length
* 92370 92470: gap of unknown length
* 92471 94334: contig of 1965 bp in length
* 94335 96676: contig of 2142 bp in length
* 96677 96776: gap of unknown length
* 96777 98217: contig of 1441 bp in length
* 98218 98317: gap of unknown length
* 98318 101563: contig of 3246 bp in length
* 101564 101664: gap of unknown length
* 101665 108765: contig of 7102 bp in length.
FEATURES
source
1..108765
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8p"
/clone="RPI-158P13"
BASE COUNT 28954 a 25034 c 24818 g 28459 t 1500 others
ORIGIN
Query Match 26.5%; Score 105; DB 2; Length 108765;
Best Local Similarity 66.7%; Pred. No. 3.8e-23;
Matches 150; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
Qy 1 atcagagctacagctgaccgagctgcccgcctatcgcaatgtgttgaggccctctgg 60
Db 68859 ATGCAGAGTTTGAATCCAGATCCCAAGCCGATGATCAAGCATCTAGCTAGGAGCCCTCAAG 68918
Qy 61 aggtattataagacgaggccctatgagccctatgagggcctgaaacgctcacagaaca 120

Db 68919 AAAATCATCGGACGAGGCTCTGGAGGCCCTTGGAGGCGTCAACGTCATGATCATG 68978
Qy 121 ggcgcaggccctggccacgccctttatttgcctgctacgaaaagttaaaaaagacattg 180
Db 68979 GGTGAGGGCCAGCCATGCTATTTTGGCTGCTATGAAAACATGAAAAGGACTTTA 69038
Qy 181 agtgaatgaatccacccctgggggcaatagccatattgccaatggt 225
Db 69039 AATGACGTTTCCACGACCAAGAAACAGCCACCTAGCCACGGT 69083

RESULT 15
AX306258
LOCUS AX306258 483 bp DNA linear PAT 11-DEC-2001
DEFINITION Sequence 1009 from Patent WO0188188.
ACCESSION AX306258
VERSION AX306258.1 GI:17645513
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (sites)
REFERENCE
AUTHORS Ishikawa,K., Asai,S., Takahashi,Y., Nagata,T. and Ishii,Y.
TITLE Method for examining ischemic conditions
JOURNAL Patent: WO 0188188-A 1009 22-NOV-2001;
School Juridical Person Nihon University (JP)
FEATURES
Location/Qualifiers
source
1. 483
/organism="Mus musculus"
/db_xref="taxon:10090"

BASE COUNT 127 a 139 c 121 g 96 t
ORIGIN

Query Match 26.5%; Score 104.8; DB 6; Length 483;
Best Local Similarity 94.0%; Pred. No. 1.8e-23;
Matches 109; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 165 gttaaaaaagacattgagtgatgtaataccacctgggggcaatagccatattgccaatgg 224
Db 15 GTTAAAGAGACATTGAGTGAGCTAATCCACCCAGGGGGCAATAGCCATATGCCAATGG 74
Qy 225 tgcggccgggtgtgtgccaattacttcattgatgcagccatgaaccctgcggaag 280
Db 75 TGCAGCCGGATGTGTGGCGACATTACTTCATGATGCAGCCATGAATCCAGCGGAAG 130

Search completed: September 28, 2002, 08:00:04
Job time: 14925 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000. Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2002, 08:06:28 ; Search time 664.31 Seconds
(without alignments)
1023.465 Million cell updates/sec

Title: US-09-870-113-11
Perfect score: 396
Sequence: 1 atgcagagtacagctga.....ttacctgtttaccactag 396

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_032802.*

- 1: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
- 2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
- 3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
- 4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
- 5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
- 6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
- 7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
- 8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
- 9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
- 10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
- 11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
- 12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
- 13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
- 14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
- 15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
- 16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
- 17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
- 18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
- 19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
- 20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
- 21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
- 22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
- 23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
- 24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	281	71.0	1418	21 AAC74843	Human ORF398
2	280	70.7	1294	22 AAI60661	Human polynucleoti
3	280	70.7	1322	21 AAF27733	Human transport pr
4	280	70.7	1336	21 AAC90457	Human uncoupling p
5	273	68.9	401	22 AAF66156	Novel human polynu
6	272	68.7	1316	22 AAI58875	Human polynucleoti
7	225	56.8	2502	22 AAS03906	Human secreted pro
8	225	56.8	27960	22 AAK69779	Human immune/haema
9	225	56.8	27960	22 AAK73320	Human immune/haema

10	174	43.9	339	22	AAK58672	Human immune/haema
11	129.6	32.7	1902	22	AAK52172	Human polynucleoti
12	128	32.3	1305	22	AAF59920	Human mitochondria
13	128	32.3	1716	21	AAK77189	Human ORF2744
14	128	32.3	2562	22	AAH90057	Human bone marrow
15	121	30.6	300	20	AAZ13026	Human gene express
16	121	30.6	710	20	AAZ15876	Human bone marrow
17	107.6	27.2	1913	22	AAH89944	Human gene express
18	105	26.5	842	22	AAK92357	Human cDNA 5'-end
19	105	26.5	842	22	AAK93895	Human cDNA clone r
20	105	26.5	1814	22	AAK94871	Human full-length
21	105	26.5	2837	22	AAK80493	Human immune/haema
22	104.8	26.5	483	24	ABI99871	Mouse ischaemic co
23	103.4	26.1	1890	22	AAK94897	Human full-length
24	103.4	26.1	3028	22	AAK91320	Human polynucleoti
25	101.8	25.7	1468	21	AAK77175	Human ORF2730
26	101	25.5	622	22	AAK91764	Human cDNA 5'-end
27	101	25.5	622	22	AAK93200	Human cDNA clone r
28	91.4	23.1	1758	20	AAV84573	Human secreted pro
29	91.4	23.1	1758	22	ABA83356	Human secreted pro
30	58	14.6	2037	24	AAS62787	cDNA sequence #574
31	42.2	10.7	2137	23	ABL08627	Drosophila melanog
32	37.4	9.4	8211	22	AAK80490	Human immune/haema
33	35	8.8	3842	22	AAK44972	cDNA encoding nove
34	34.4	8.7	6712	21	AAK76691	Human ORF2246
35	34	8.6	673	21	AAF12614	Aspergillus oryzae
36	33.6	8.5	1371	21	AAK47438	Arabidopsis thalia
37	33.6	8.5	1373	21	AAK39502	Arabidopsis thalia
38	33.6	8.5	11800	22	AAK82152	Human immune/haema
39	33.2	8.4	1227	23	AAK81309	DNA encoding novel
40	32.4	8.2	1311	18	AAK68957	Blackcurrant prir7
41	32.4	8.2	5150	18	AAK68952	Blackcurrant fruit
42	32.4	8.2	6326	23	ABL08626	Drosophila melanog
43	32	8.1	355	22	AAK26927	Streptococcus muta
44	32	8.1	1114	24	AAI71019	Human mitochondria
45	32	8.1	1363	20	AAK60283	cDNA encoding a pr

ALIGNMENTS

RESULT 1
AAC74843
ID AAC74843 standard; cDNA; 1418 BP.
XX
AC AAC74843;
XX
DT 08-FEB-2001 (first entry)
XX
XX Human ORF398 polynucleotide sequence SEQ ID NO:795.

Human; open reading frame; ORF3; detection; cytostatic; hepatotropic;
vulnary; antipsoriatic; antiparkinsonian; neurotropic; neuroprotective;
anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiac;
immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
hypotensive; dermatological; immunosuppressive; antiinflammatory;
antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
neurodegenerative disorder; osteoarthritis; graft vs host disease;
cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
cholesterol ester storage; systemic lupus erythematosus; infection;
severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
bone damage; cartilage damage; antiinflammatory disease; coagulation;
thrombosis; contraceptive; ss.

Homo sapiens.

WO200058473-A2.

05-OCT-2000.

31-MAR-2000; 2000WO-US08621.

```
XX 31-MAR-1999; 99US-0127607.
PR 02-APR-1999; 99US-0127636.
PR 05-APR-1999; 99US-0127728.
PR 30-MAR-2000; 2000US-0540763.
XX (CURA-) CURAGEN CORP.
XX ShImkets RA, Leach M;
XX
XX WPI; 2000-602362/57.
DR P-PSDB; AAB40634.
XX
XX Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -
XX
XX Claim 5; Page 853-854; 5507pp; English.
XX
XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human OREF open reading frames 1 to 3161. The OREF
CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
CC antiproliferative; antiparkinsonian; neurotropic; neuroprotective;
CC osteoprotic; anticonvulsant; antiarthritic; immunosuppressant;
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC antithyroid; and antianemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an OREF-associated disorder. The
CC nucleic acids can be used to express OREF proteins in gene therapy
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.
XX
XX Sequence 1418 BP; 353 A; 400 C; 368 G; 297 T; 0 other;
SQ

Query Match 71.0%; Score 281; DB 21; Length 1418;
Best Local Similarity 100.0%; Pred. No. 2.9e-85;
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atgcagagctcacagctgaccagctgcccgcgtatcgcaatgtgttgaggccctctgg 60
Db 178 atgcagagctcacagctgaccagctgcccgcgtatcgcaatgtgttgaggccctctgg 237
Qy 61 aggtattataaagcaggggcctatgaggccatgaggggcgtgacgtcacagcaaca 120
Db 238 aggtattataaagcaggggcctatgaggccatgaggggcgtgacgtcacagcaaca 297
Qy 121 ggcgcaggccctgcccagcccttatttgcctgctacgaaaaagttaaaaaagacattg 180
Db 298 ggcgcaggccctgcccagcccttatttgcctgctacgaaaaagttaaaaaagacattg 357
Qy 181 agtgatgtaattccaccctgggggccaatagccatattgccaatggtgcggcggtgtgtg 240
Db 358 agtgatgtaattccaccctgggggccaatagccatattgccaatggtgcggcggtgtgtg 417
Qy 241 gcaacattacttcatgatgcagccatgaacccctgcggaag 281
Db 418 gcaacattacttcatgatgcagccatgaacccctgcggaag 458

RESULT 2
AAI60661/C
ID AAI60661 standard; cDNA; 1294 BP.
XX
```

```
AC AAI60661;
XX
XX 22-OCT-2001 (first entry)
DE Human polynucleotide SEQ ID NO 4650.
XX
XX Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200153312-A1.
PN
XX
XX 26-JUL-2001.
PD
XX
XX 26-DEC-2000; 2000WO-US34263.
PF
XX
XX 21-JAN-2000; 2000US-0488725.
PR
XX 25-APR-2000; 2000US-0552317.
PR
XX 09-JUL-2000; 2000US-0598042.
PR
XX 19-JUL-2000; 2000US-0620312.
PR
XX 03-AUG-2000; 2000US-0653450.
PR
XX 14-SEP-2000; 2000US-0662191.
PR
XX 19-OCT-2000; 2000US-0693036.
PR
XX 29-NOV-2000; 2000US-0727344.
XX
XX (HYSE-) HYSEQ INC.
PA
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang V, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
XX WPI; 2001-442253/47.
DR
XX P-PSDB; AAM41505.
DR
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
PT
XX
XX Claim 1; SEQ ID NO 4650; 10078pp; English.
PS
XX
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nontropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
XX Sequence 1294 BP; 267 A; 335 C; 368 G; 324 T; 0 other;
SQ

Query Match 70.7%; Score 280; DB 22; Length 1294;
Best Local Similarity 100.0%; Pred. No. 6.1e-85;
Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atgcagagctcacagctgaccagctgcccgcgtatgcaatgtgttgaggccctctgg 60
Db 1134 ATGCAGAGCTCAGAGCTGACCCAGCTGCCCGCTATCGCATGTGTGGAGGCCCTCTGG 1075
Qy 61 aggtattataaagcaggggcctatgaggccatgaggggcgtgacgtcacagcaaca 120
```



```
Db 1074 AGGATTATAAGAACGAGGCGCTATGGAGGCCCATGAGGGGCTGAACGTACAGCAACA 1015
QY 121 ggcgcaggcccgccacccctttatttgcctgctacgaaagttaaaaaagacattg 180
Db 1014 GCGCAGGCGCTGCCACGCCCTTTATTTTGGCTGTGCTAGCAAAAGTTAAAGACATTG 955
QY 181 agtcatgtaataccacccctgggggcaatagcccatattgccaatggcgccgggtgtg 240
Db 954 AGTGATGTAATCCACCTGGGCAATAGCCCATATTGCCAATGGTCGCGCGGGTGTGTG 895
QY 241 gcaacattactcatgatcagccatgaaccctgcggaag 280
Db 894 GCAACATTACTTCATGATCAGCCATGAACCCCTGGCGGAAG 855

RESULT 3
AAF27733
ID AAF27733 standard; cDNA; 1322 BP.
XX AC AAF27733;
XX DT 28-MAR-2001 (first entry)
XX DE Human transport protein TPPT-33 coding sequence.
XX KW Human; transport protein; TPPT; transport disorder; metabolic disorder;
KW neurological disorder; cardiovascular disorder; reproductive disorder;
KW immune disorder; cancer; ss.
XX OS Homo sapiens.
XX PN WO200078953-A2.
XX PD 28-DEC-2000.
XX PF 16-JUN-2000; 2000WO-US16668.
XX PR 17-JUN-1999; 99US-0139923.
XX PR 10-AUG-1999; 99US-0148177.
XX PR 18-AUG-1999; 99US-0149357.
XX PR 28-OCT-1999; 99US-0162287.
XX PA (INCY-) INCYTE GENOMICS INC.
XX PI Lal P, Yang J, Yue H, Hillman JL, Tang YT, Bandman O, Burford N;
PI Baughn MR, Azimzai Y, Lu DAM, Au-Young J, Patterson C;
XX WPI: 2001-041424/05.
XX DR P-PSDB; AAB60113.
XX Isolated polypeptide with a human transport protein sequence is useful
PT for the diagnosis, prevention and treatment of disorders associated
PT with the immune, reproductive and cardiovascular systems -
XX-
PS Claim 5; Page 160; 165pp; English.
XX The present invention provides the protein and coding sequences for 43
CC novel human transport proteins (designated TPPTs). These can be used in
CC the diagnosis and treatment of transport, metabolic, neurological,
CC reproductive, cardiovascular and immune disorders, and cell proliferative
CC disorders such as cancer.
XX Sequence 1322 BP; 328 A; 376 C; 351 G; 267 T; 0 other;

Query Match 70.7%; Score 280; DB 22; Length 1322;
Best Local Similarity 100.0%; Pred. No. 6.2e-85;
Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgcagagtctacagctgacccagctgccgctatcgcaatgtgttgaggccctctgg 60
Db 186 atgcagagtctacagcctgacccagctgccgctatcgcaatgtgttgaggccctctgg 245
```

```
QY 61 aggattataaagcagagggccctatgagggcccatgagggcctgaacgtcacagcaaca 120
Db 246 aggattataaagcagagggccctatgagggcccatgagggcctgaacgtcacagcaaca 305
QY 121 ggcgcaggccctgcccacccctttatttgcctgctacgaaagttaaaaaagacattg 180
Db 306 ggcgcaggccctgcccacccctttatttgcctgctacgaaagttaaaaaagacattg 365
QY 181 agtcatgtaataccacccctgggggcaatagcccatattgccaatggcgccgggtgtg 240
Db 366 agtcatgtaataccacccctgggggcaatagcccatattgccaatggcgccgggtgtg 425
QY 241 gcaacattactcatgatcagccatgaaccctgcggaag 280
Db 426 gcaacattactcatgatcagccatgaaccctgcggaag 465

RESULT 4
AAC90457
ID AAC90457 standard; cDNA; 1336 BP.
XX AC AAC90457;
XX DT 12-MAR-2001 (first entry)
XX DE Human uncoupling protein cDNA #6.
XX KW Human; uncoupling protein; immunosuppressive; antiarthritic;
KW antirheumatic; antiproliferative; cardiant; vasotropic;
KW cerebroprotective; neuroprotective, antibacterial; ophthalmological;
KW gastrointestinal; nephrotropic; gynaecological; vulnery; thrombolytic;
KW gene therapy; cancer; wound; infectious disease; thrombosis; arthritis;
KW infertility; ss.
XX OS Homo sapiens.
XX PN WO200061614-A2.
XX PD 19-OCT-2000.
XX PF 06-APR-2000; 2000WO-US09534.
XX PR 09-APR-1999; 99US-0128701.
XX PR 08-JUL-1999; 99US-0142821.
XX PR 18-AUG-1999; 99US-0149448.
XX PR 12-NOV-1999; 99US-0164751.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Ruben SM, Ni J, Komatsoulis G, Rosen CA, Soppet DR;
XX WPI: 2000-656322/63.
XX DR P-PSDB; AAB50383.
XX Uncoupling proteins and nucleic acid sequences encoding them, useful
PT for detecting, preventing and treating proliferative, neurological,
PT immune system, cardiovascular and gastrointestinal disorders -
XX-
PS Claim 1; Page 306-307; 343pp; English.
XX The present sequence is one of eighteen isolated nucleotide sequences
CC encoding uncoupling proteins. The nucleotide sequences may be used for
CC the detection of various disorders such as cancer, for chromosome
CC identification, as chromosome markers and for numerous other diagnostic
CC or research purposes. The uncoupling protein encoded by the nucleotide
CC sequences may be used to treat disorders such as neural, immune,
CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
CC renal and proliferative disorders, wounds, infectious diseases,
CC thrombosis, arthritis, and infertility.
XX Sequence 1336 BP; 333 A; 377 C; 358 G; 268 T; 0 other;
```

```
Query Match      70.7%; Score 280; DB 21; Length 1336;
Best Local Similarity 100.0%; Pred. No. 6.2e-85;
Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atgcagagtctacagcctgacccagctcccgctatcgcaatgtgttgaggccctctgg 60
    |||
Db 197 atgcagagtctacagcctgacccagctcccgctatcgcaatgtgttgaggccctctgg 256

Qy 61 aggattataagaacgaggcctatggagcccatgagggcctgaacgtcacagcaaca 120
    |||
Db 237 aggattataagaacgaggcctatggagcccatgagggcctgaacgtcacagcaaca 316

Qy 121 ggcgcaggcctgcccagcccttatttgcctgctacgaaagttaaaaaagacattg 180
    |||
Db 317 ggcgcaggcctgcccagcccttatttgcctgctacgaaagttaaaaaagacattg 376

Qy 181 agtgaatgaatccaccctgggggcaatagccatattgccaatgtgcggcgggtgtgtg 240
    |||
Db 377 agtgaatgaatccaccctgggggcaatagccatattgccaatgtgcggcgggtgtgtg 436

Qy 241 gcaacattactctatgatgcagccatgaacctgcggaag 280
    |||
Db 437 gcaacattactctatgatgcagccatgaacctgcggaag 476

RESULT 5
AAF66156
ID AAF66156 standard; cDNA; 401 BP.
XX
AC AAF66156;
XX
DT 09-APR-2001 (first entry)
XX
DE Novel human polynucleotide, SEQ ID NO: 1912.
XX
KW Human; cytostatic; gene therapy; colon cancer; prostate cancer;
KW breast cancer; lung cancer; cancer detection; ss.
XX
OS Homo sapiens.
XX
PN WO200102568-A2.
XX
PD 11-JAN-2001.
XX
PF 30-JUN-2000; 2000WO-US18374.
XX
PR 02-JUL-1999; 99US-0142310.
XX
PR 02-JUL-1999; 99US-0142311.
XX
PA (CHIR ) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
PI Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A;
PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G, Drmanac R;
PI KrakenJakov R, Drmanac S, Dickson M, Labat I, Leshkowitz D;
PI Kita D, Garcia V, Jones LW, Strache-Crain B;
XX
DR WPI; 2001-091805/10.
XX
PT Library of polynucleotides for diagnosing a cancerous state of a
PT mammalian cell and detecting cancer, particularly of the colon or
PT prostate, comprises 3351 human polynucleotide sequences -
XX
PS Claim 9; Page 818; 1046pp; English.
XX
CC The present sequence is one of 3351 sequences in a library of human
CC polynucleotides. The library is used to detect differentially expressed
CC genes correlated with a cancerous state of a mammalian cell and can
CC detect colon, prostate, breast and lung cancer. The library can be used
CC to produce probes for detection of mRNA and to produce additional copies
CC of the polynucleotides. The probes can be used for chromosome mapping of
CC the polynucleotide and for detection of transcription levels. Ribozymes
```

```
CC or antisense oligonucleotides can be generated. The polynucleotides and
CC their gene products are used as genetic or biochemical markers (e.g. in
CC blood or tissues) that will detect the earliest changes along the
CC carcinogenesis pathway and/or monitor the efficacy of therapies and
CC preventive interventions. The polynucleotides, polypeptides and
CC antibodies against them can be used in pharmaceutical compositions to
CC treat the cancers and proliferative disorders such as neoplasia,
CC dysplasia and hyperplasia.
XX
SQ Sequence 401 BP; 100 A; 104 C; 120 G; 76 T; 1 other;

Query Match      68.9%; Score 273; DB 22; Length 401;
Best Local Similarity 100.0%; Pred. No. 8.2e-83;
Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 gtctacagcctgacccagctcccgctatcgcaatgtgttgaggccctctggaggatta 67
    |||
Db 9 gtctacagcctgacccagctcccgctatcgcaatgtgttgaggccctctggaggatta 68

Qy 68 taagaacgaggggcctatggagggcccatgaggggctgaacgtcacagcagcgcgag 127
    |||
Db 69 taagaacgaggggcctatggagggcccatgaggggctgaacgtcacagcagcgcgag 128

Qy 128 ggcctgcccacgccccttatttgcctgctacgaaaagttaaaaaagacattgagtgtg 187
    |||
Db 129 ggcctgcccacgccccttatttgcctgctacgaaaagttaaaaaagacattgagtgtg 188

Qy 188 taatccaccctgggggcaatagccaatttgccaatgtgcggcgggtgtgtggcaacat 247
    |||
Db 189 taatccaccctgggggcaatagccaatttgccaatgtgcggcgggtgtgtggcaacat 248

Qy 248 tacttcatgatgcagccatgaacctgcggaag 280
    |||
Db 249 tacttcatgatgcagccatgaacctgcggaag 281

RESULT 6
AAI58875
ID AAI58875 standard; cDNA; 1316 BP.
XX
AC AAI58875;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 1078.
XX
KW Human; neurotropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
XX
PR 25-APR-2000; 2000US-0552317.
XX
PR 09-JUL-2000; 2000US-0598042.
XX
PR 19-JUL-2000; 2000US-0620312.
XX
PR 03-AUG-2000; 2000US-0653450.
XX
PR 14-SEP-2000; 2000US-0662191.
XX
PR 19-OCT-2000; 2000US-0693036.
XX
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
```


XX DT 06-NOV-2001 (first entry)
XX DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:24591.
XX KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytotstatic; gene therapy; vaccine; metastasis; ds.
XX OS Homo sapiens.
XX PN WO200157182-A2.
XX PD 09-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US01354.
XX PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.

PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-483426/52.
DR
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis
XX
XX Disclosure; SEQ ID NO 24591; 3071pp + Sequence Listing; English.
PS
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent, the
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 27960 BP; 7998 A; 5835 C; 6583 G; 7544 T; 0 other;

Query Match 56.8%; Score 225; DB 22; Length 27960;
Best Local Similarity 100.0%; Pred. No. 1.7e-65;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atgcagagtctacagcctgaccagctgcccgcctatcgcaatgtgttgaggccctctgg 60
|||||
Db 18295 atgcagagtctacagcctgaccagctgcccgcctatcgcaatgtgttgaggccctctgg 18354
|||||

Qy 61 aggattataagaacagagggccctatggagcccatgaggggctgaacgtcacagcaaca 120
|||||
Db 18355 aggattataagaacagagggccctatggagcccatgaggggctgaacgtcacagcaaca 18414
|||||

Qy 121 ggcgcagggcctgcccacgcctctatttgcctctacgaaaagttaaaaaagacattg 180
|||||
Db 18415 ggcgcagggcctgcccacgcctctatttgcctctacgaaaagttaaaaaagacattg 18474
|||||

Qy 181 agtgatgtaataccacctgggggcaatgccaatattgccaatggt 225
|||||
Db 18475 agtgatgtaataccacctgggggcaatgccaatattgccaatggt 18519
|||||

RESULT 9
AAK73320
ID AAK73320 standard; DNA; 27960 BP.
XX
AC AAK73320;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:28132.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.

XX Homo sapiens.
OS WO200157182-A2.
PN
XX 09-AUG-2001.
PD
XX
PF 17-JAN-2001; 2001WO-US01354..
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 21-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.

PF 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.

PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
DR P-PSDB; AAM85891.
XX

XX Human mitochondrial solute carrier protein; hMSC-o; hypothalamus;
 KW preparation; detection; ss.
 XX Homo sapiens.
 OS CN1269409-A.
 XX 11-OCT-2000.
 XX 17-MAR-2000; 2000CN-0114958.
 XX 17-MAR-2000; 2000CN-0114958.
 XX (SREN-) SOUTHERN RES CENT NAT HUMAN GENE GROUP.
 XX Zhang X, Gao X, Xiao H;
 XX WPI; 2001-050544/07.
 DR P-PSDB; AAB60658.
 XX New human mitochondrion solute carrier protein and its nucleic acid -
 XX Claim 1; Page 19-20; 21pp; Chinese.
 CC The invention relates to a novel human mitochondrial solute carrier
 CC protein, hMSC-o (AAB60658), and cDNA encoding it (AAF59920). hMSC-o is
 CC expressed in normal human hypothalamus tissue. The invention also relates
 CC to the preparation of hMSC-o proteins and nucleic acids, and the
 CC detection of hMSC-o proteins and nucleic acids in a sample. The present
 CC sequence represents cDNA encoding hMSC-o.
 XX
 SQ Sequence 1305 BP; 292 A; 388 C; 363 G; 262 T; 0 other;
 Query Match 32.3%; Score 128; DB 22; Length 1305;
 Best Local Similarity 66.1%; Pred. No. 4.5e-33;
 Matches 185; Conservative 0; Mismatches 95; Indels 0; Gaps 0;
 QY 1 atgcagagctctacagctgaccagctgcccgtctatcgcaatgtgtgagccctctg 60
 Db 172 atgcagagtttgatgcagatcccaagcccgatcacagaagtatctacgagccctcaag 231
 QY 61 aggattataagaacgagggccctatggagcccatgaggggctgaacgtcacagcaaca 120
 Db 232 aaaaatcgagcgaagctctgagggcccttgagggcgtcaacgtcatgacttg 291
 QY 121 ggcgagggcctgcccagccctttatttctgctgtctacgaaagttaaaaagacattg 180
 Db 292 ggtcgagggccgcccattgctgtatttctgctgtatgctgtatgaaacatgaaaggactta 351
 QY 181 agtgaatgaatccaccctggggccatagccatattgccaatggtgcccgggtgtgtg 240
 Db 352 atagcgttttccaccacacgaagaaacacccacccatgacgaaggatagctgggagatg 411
 QY 241 gcaacattactctatgatgcagccatgaacctggggaag 280
 Db 412 gccacctgctcagatcggtgtaataatgaatccagcagaag 451
 RESULT 13
 AAC77189
 ID AAC77189 standard; cDNA; 1716 BP.
 AC AAC77189;
 XX
 DT 08-FEB-2001 (first entry)
 XX Human ORFX ORF2744 polynucleotide sequence SEQ ID NO:5487.
 DE Human; open reading frame; ORFX; detection; cytosolic; hepatotropic;
 KW vulnary; antiparotiatric; antiparkinsonian; nootropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiac;

KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antinflammatory;
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antinflammatory disease; coagulation;
 KW thrombosis; contraceptive; ss.
 XX Homo sapiens.
 XX WO200058473-A2.
 XX 05-OCT-2000.
 XX 31-MAR-2000; 2000WO-US08621.
 XX 31-MAR-1999; 99US-0127607.
 PR 02-APR-1999; 99US-0127636.
 PR 05-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.
 XX
 XX (CURA-) CURAGEN CORP.
 PA Shinkets RA, Leach M;
 XX WPI; 2000-602362/57.
 DR P-PSDB; AAB42980.
 XX
 PT Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX
 XX Claim 5; Page 4661-4662; 5507pp; English.
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnary;
 CC antiparotiatric; antiparkinsonian; nootropic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
 CC immunostimulant; cardiac; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; antirheumatic;
 CC antinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
 CC antithyroid; and antianemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.
 XX
 SQ Sequence 1716 BP; 411 A; 457 C; 446 G; 402 T; 0 other;
 Query Match 32.3%; Score 128; DB 21; Length 1716;
 Best Local Similarity 66.1%; Pred. No. 5.2e-33;
 Matches 185; Conservative 0; Mismatches 95; Indels 0; Gaps 0;
 QY 1 atgcagagctctacagctgaccagctgcccgtctatcgcaatgtgtgagccctctg 60
 Db 256 atgcagagtttgatgcagatcccaagcccgatcacagaagtatctacgagccctcaag 315
 QY 61 aggattataagaacgaggggctctatggagcccatgaggggctgaacgtcacagcaaca 120

Db 316 aaatcatcgagaccggaagctttctgagcccttgagcccttgagcgctcaacgctcatgcatg 375
 Qy 121 ggcagagcctccacgcccctttatttgcctgtctacgaaagttaaaaaagacattg 180
 Db 376 ggtgagggcagcccatgcatgtatttgcctgtcatgaaacatgaaagagacttta 435
 Qy 181 agtgatgaataccacctgggggcaatagccatatttcccaatggtgcgcgggtgtgtg 240
 Db 436 aatgacgttttccaccaccaaagaaacagccacctagccaacgggatagctgggagtagt 495
 Qy 241 gcaacattactcatgatgacgagccatgaacccctgcggaag 280
 Db 496 gccacctgtcccagatgcggtgaatgaatccagcagaag 535

 RESULT 14
 AAH90057
 ID AAH90057 standard; cDNA; 2562 BP.
 AC AAH90057;
 XX
 XX
 DT 01-OCT-2001 (first entry)
 XX
 XX Human bone marrow cDNA, SEQ ID NO: 301.
 DE
 XX
 XX Human; bone marrow; antiinflammatory; cytostatic; neuroprotective;
 KW antiviral; antibacterial; antifungal; anti-HIV; haemostatic;
 KW immunosuppressive; gene therapy; cytokine cell proliferation;
 KW cell differentiation modulator; immune disorder; infection; cancer;
 KW human immunodeficiency virus; HIV; autoimmune disorder; haemophilia; ss.
 XX
 OS Homo sapiens.
 XX
 XX WO200153453-A2.
 XX
 XX 26-JUL-2001.
 XX
 XX 23-DEC-2000; 2000WO-US34960.
 XX
 XX 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 30-NOV-2000; 2000US-0250583.
 XX
 XX (HYSE-) HYSEQ INC.
 XX
 XX Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;
 PI Ren F, Wang J, Werhman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Drmanac RT;
 XX
 DR WPI; 2001-488707/53.
 DR P-PSDB; AAM00938.
 XX
 XX Novel bone-marrow-expressed polynucleotides and polypeptides, useful
 PT for treating e.g. cancer and immune deficiency disorders -
 XX
 XX Claim 1; Page 412-413; 648pp; English.
 PS
 XX
 XX The present sequence is one of 251 novel human polynucleotides
 CC expressed in the bone marrow. The polynucleotide and the
 CC polypeptide encoded by it are useful in the treatment of various
 CC immune deficiencies and disorders. The deficiencies and disorders may
 CC be genetic, or may be caused by a viral (e.g. HIV), bacterial or fungal
 CC infection, or may result from an autoimmune disorder, a coagulation
 CC disorder (e.g. haemophilia), inhibition of tumour cell proliferation,
 CC suppression of an inflammatory response or treatment of a nervous
 CC system disorder such as Alzheimer's disease. Detection of the presence
 CC or increased expression of the polynucleotide or the protein it
 CC encodes is useful for the diagnosis and/or prognosis of one

CC or more types of cancer. The polynucleotide and polypeptide can be
 CC used as nutritional sources or supplements and in the screening of
 CC chemical compounds as potential drugs.
 XX
 SQ Sequence 2562 BP; 623 A; 660 C; 627 G; 652 T; 0 other;

 Query Match 32.3%; Score 128; DB 22; Length 2562;
 Best Local Similarity 66.1%; Pred. No. 6.4e-33;
 Matches 185; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

 Qy 1 atgcagagtctacagctgacccagctgccgcgtatcgcgaatgtgttgagccctctg 60
 Db 2021 atgcagagtttgatccagatcccaagcccaagtcacacagtagtctctacgagccctcaag 2080
 Qy 61 aggattataaagacgagggcctatgagcccatgaggggctggaacgtcacagcaaca 120
 Db 2081 aaatcatcgagaccgaagcttctgagcccttgagcgctcaacgtcatgcatc 2140
 Qy 121 ggcgagggcctgccacgcctttatttgcctgtctacgaaagttaaaaaagacattg 180
 Db 2141 ggtgagggcgcgcctatgcatgtatttgcctgtcatgaaacatgaaagagacttta 2200
 Qy 181 agtgatgaataccacctgggggcaatagccatatttcccaatggtgcgcgggtgtgtg 240
 Db 2201 aatgacgttttccaccaccaaagaaacagccacctagccaacgggatagctgggagtagt 2260
 Qy 241 gcaacattactcatgatgacgagccatgaacccctgcggaag 280
 Db 2261 gccacctgtcccagatgcggtgaatgaatccagcagaag 2300

 RESULT 15
 AAZ13026
 ID AAZ13026 standard; cDNA; 300 BP.
 XX
 XX AC AAZ13026;
 XX
 DT 12-OCT-1999 (first entry)
 XX
 XX Human gene expression product cDNA sequence SEQ ID NO:495.
 DE
 XX Human; gene; gene expression product; diagnosis; therapy; probe;
 KW detection; mapping; tissue typing; profiling; forensic; cancer;
 KW genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.
 XX
 OS Homo sapiens.
 XX
 XX WO9938972-A2.
 XX
 XX 05-AUG-1999.
 XX
 XX 28-JAN-1999; 99WO-US01619.
 XX
 XX 03-APR-1998; 98US-0080666.
 PR 28-JAN-1998; 98US-0072910.
 PR 24-FEB-1998; 98US-0075954.
 PR 31-MAR-1998; 98US-0080114.
 PR 03-APR-1998; 98US-0080515.
 XX
 XX (CHIR) CHIRON CORP.
 PA (HYSE-) HYSEQ INC.
 XX
 XX Crkvenjakov R, Dickson M, Drmanac R, Drmanac S;
 PI Escobedo J, Garcia PD, Garcia V, Giese K, Innis MA;
 PI Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;
 PI Lanson G, Leshkowitz D, Pot B, Randazzo F, Reinhard C;
 PI Stache-Crain B, Sudduth-Klinger J, Williams LT;
 XX
 DR WPI; 1999-494092/41.
 XX
 XX Novel human genes and their expression products which are
 PT differentially expressed in different cell types

Search completed: September 28, 2002, 08:06:37
Job time: 11592 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2002, 07:56:00 ; Search time 192.06 Seconds
(without alignments)
534.279 Million cell updates/sec

Title: US-09-870-113-11
Perfect score: 396
Sequence: 1 atgcagagctctacagcctga.....ttacctgttttaccactag 396

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
- 2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
- 3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
- 4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
- 5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	32.4	8.2	1311	4	US-09-068-140A-9
2	32.4	8.2	5150	4	US-09-068-140A-14
3	32	8.1	1643	2	US-08-933-750C-68
4	32	8.1	1643	3	US-09-234-613-68
5	30.2	7.6	2277	1	US-08-676-967-2
6	30.2	7.6	2277	1	US-08-676-974-2
7	30.2	7.6	2277	2	US-09-098-487-2
8	29.6	7.5	4091	4	US-08-630-915A-33
9	29.2	7.4	5532	2	US-08-475-035-3
10	29	7.3	1747	1	US-08-202-389-1
c 11	29	7.3	8600	4	US-09-457-708-1
c 12	28.6	7.2	1311	3	US-09-073-569-4
c 13	28.4	7.2	4368	1	US-08-026-138E-17
c 14	28.4	7.2	4446	1	US-08-026-138E-6
c 15	28.2	7.1	7011	4	US-09-268-163-9
c 16	27.8	7.0	1323	4	US-09-297-937C-10
c 17	27.8	7.0	4665	4	US-09-297-937C-12
c 18	27.8	7.0	5643	1	US-08-144-602B-4
c 19	27.6	7.0	3942	4	US-09-162-484-19
c 20	27.2	6.9	3771	1	US-08-185-432-3
c 21	27.2	6.9	3771	1	US-08-185-432-23
c 22	27.2	6.9	3810	2	US-08-475-844-8
c 23	27.2	6.9	3810	5	PCT-US95-08429-8
c 24	27.2	6.9	5063	1	US-08-185-432-1
c 25	27	6.8	591	3	US-08-686-878A-5
c 26	27	6.8	2379	3	US-08-537-361E-5
c 27	27	6.8	3153	4	US-09-175-928-9

28	27	6.8	3508	4	US-08-462-457B-7	Sequence 7, Appli
29	27	6.8	3508	4	US-08-334-179A-7	Sequence 7, Appli
30	27	6.8	7653	4	US-08-471-112A-1	Sequence 1, Appli
31	27	6.8	7824	5	PCT-US95-06722-11	Sequence 11, Appl
c 32	26.8	6.8	1128	3	US-09-106-217-15	Sequence 15, Appl
c 33	26.8	6.8	1134	3	US-09-106-217-1	Sequence 1, Appli
34	26.8	6.8	1192	4	US-09-142-565-1	Sequence 1, Appli
35	26.8	6.8	1255	1	US-08-518-878B-38	Sequence 38, Appl
36	26.8	6.8	1255	1	US-08-294-522B-38	Sequence 38, Appl
37	26.8	6.8	1255	2	US-08-470-868A-38	Sequence 38, Appl
38	26.8	6.8	1596	2	US-08-807-861A-38	Sequence 38, Appl
39	26.8	6.8	1596	3	US-09-210-681-38	Sequence 38, Appl
40	26.8	6.8	1596	3	US-08-946-719A-38	Sequence 38, Appl
c 41	26.8	6.8	3672	4	US-09-056-105-17	Sequence 17, Appl
42	26.8	6.8	246240	2	US-08-724-394A-20	Sequence 20, Appl
43	26.8	6.8	246240	2	US-08-724-394A-21	Sequence 21, Appl
44	26.8	6.8	246240	2	US-08-724-394A-22	Sequence 22, Appl
45	26.6	6.7	573	2	US-08-290-665A-138	Sequence 138, App

ALIGNMENTS

RESULT 1
US-09-068-140A-9
; Sequence 9, Application US/09068140A
; Patent No. 6281409
; GENERAL INFORMATION:
; APPLICANT: Mary Rose Woodhead, Mark Andrew Taylor
; APPLICANT: and Rex Michael Brennan
; TITLE OF INVENTION: Blackcurrant Promoters and Genes
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/068,140A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP96/04807
; FILING DATE: No. 6281409ember 4, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Dinner, Dara L.
; REGISTRATION NUMBER: 33,680
; REFERENCE/DOCKET NUMBER: C70237
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5017
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1311 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Ribes nigrum
; STRAIN: Ben Alder
US-09-068-140A-9

Db 482 CTACCCGCTTCTACTGGCACCATTGGATGCCTTCGTGAAGATCTGTGAGGCACGAGGGGACCA 541
QY 86 ggagggcccatgaggggctgaacgtcacagcaacaggcgagggcctgccccacgccccttt 145
Db 542 GGACCCCTCTGGAGCGGCTCCCGCCACACCCTGGTGATGACTGTGCAGCTACCGCCATCT 601
QY 146 atttgcctgtctacgaaagttaaaaaagacattgagtgatgataatcc 193
Db 602 ACTTCACTGCCTATGACCAACTGAAGGCTTCCTGTGTGGTCGAGCCC 649

RESULT 4
US-09-234-613-68
; Sequence 68, Application US/09234613
; Patent No. 6132973
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/234,613
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,750
; FILING DATE: September 23, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0356 US
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1643 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: SINO00T01
; CLONE: 724157
US-09-234-613-68

Query Match 8.1%; Score 32; DB 3; Length 1643;
Best Local Similarity 49.4%; Pred. No. 0.29;
Matches 83; Conservative 0; Mismatches 85; Indels 0; Gaps 0;
QY 26 ctgcccgctatgcgaatgtgttgagggccctctggaggtatataagaacgagggccctat 85
Db 482 CTACCCGCTTCTACTGGCACCATTGGATGCCTTCGTGAAGATCTGTGAGGCACGAGGGGACCA 541

QY 86 ggagggcccatgaggggctgaacgtcacagcaacaggcgagggcctgccccacgccccttt 145
Db 542 GGACCCCTCTGGAGCGGCTCCCGCCACACCCTGGTGATGACTGTGCAGCTACCGCCATCT 601
QY 146 atttgcctgtctacgaaagttaaaaaagacattgagtgatgataatcc 193
Db 602 ACTTCACTGCCTATGACCAACTGAAGGCTTCCTGTGTGGTCGAGCCC 649
RESULT 5
US-08-676-967-2
; Sequence 2, Application US/08676967
; Patent No. 5747317
; GENERAL INFORMATION:
; APPLICANT: COLLINS, KATHLEEN
; TITLE OF INVENTION: Human Telomerase
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Science & Technology Law Group
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/676,967
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman Ph.D., Richard A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UCB96-055
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)343-4341
; TELEFAX: (415)343-4342
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2277 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-676-967-2

Query Match 7.6%; Score 30.2; DB 1; Length 2277;
Best Local Similarity 27.4%; Pred. No. 1.5;
Matches 80; Conservative 50; Mismatches 162; Indels 0; Gaps 0;
QY 86 ggagggcccatgaggggctgaacgtcacagcaacaggcgagggcctgccccacgccccttt 145
Db 503 GNAARGCNYTNAARGNATGAAYATGAARGATHAARGNMGNCNGTNGCNGTNGAYT 562
QY 146 atttgcctgtctacgaaagttaaaaaagacattgagtgatgataatccacctgggggca 205
Db 563 GGGCNGTNGCNAARGAAYARTAYAAARGAYACNCARMSNGCNATNGNGNGARGARA 622
QY 206 atagccatattgccaatgttgcggcggtgtgtgtggcaacattcttcatgatgcagcca 265
Db 623 ARWSNCAYCARWSNAARCAAYCARGARWSNGTNAARAARAARGNMGNCARGARGAYA 682
QY 266 tgaacctgcggaagttaattctctcaacctatcaactctgtgtggcagctgcacctgtga 325
Db 683 TGGARGARGARAAAYGAYGAYGAYGAYGAYGAYGAYGAYGAYGAYGAYGAYGAYGAYGAY 742
QY 326 ttcttttcagttgtcagaagaacacacatcagtttttgggtgggaaattctgt 377
Db 743 AYGAYGARGARGARGARGARAAAYATHGARWSNAARGTNAACNACNCGT 794

Query Match	7.6%	Score 30.2;	DB 1;	Length 2277;
Best Local Similarity	27.4%;	Pred. No.1.5;		
Matches	80;	Conservative 50;	Mismatches 162;	Indels 0; Gaps 0;

QY	86	ggaagcccatgaggggctgaacgtcacagcaaacagcgcgagcgcctgtgccaccgcccttt	145
		: : : : : : : : : : : : :	
Db	503	GNAARGCNYTNAARGGNATGAAYTAAGARGARYTHAARGMNMGNACNGTNGCNTNGAYT	562
QY	146	atttgcctgtactcgaaaaagttaaaaagacatttgatgatgaatccacctggggcca	205
		: : : : : :	
Db	563	GGCGCTGTCWAARGAVARTAYAAARGAYACNCARWSNGTNWWSGCCNATHGGNGARGARA	622
QY	206	atagccattatgccaatgtgtgcggcgggtgtgtggcaacattactcatgatcgacca	265
		: : : : : : : : : : : : :	
Db	623	ARWSCAYGARWSNAARCAYCARGRWSNGTNAARAAPAAARGNMGNGARGARGAYA	682
		: : : : : : : : : : : : :	
QY	266	tgaacctcgcgaggtaatgatcttcctaacctatcactctgtggcagctgcacctgta	325
		: : : : : : : : : : :	
Db	683	TGGARGARGARGAARAAYGAYGAYGAYGAYGAYGAYGARGAYGGTGNTTYG	742
QY	326	ttctottacagtttgcagagaagaagcacatcagtttgtgtgggaatacttgt	377
		: : : : : : : : : : : : : :	
Db	743	AYGAYGARGAYGARGARGARGARAAYATHGARWSNAARGTNACNAARCCNGT	794

RESULT	7
US-09-098-487-2	
; Sequence 2, Application US/09098487	
; Patent No. 5917025	
; GENERAL INFORMATION:	

Query Match	7.6%	Score 30.2;	DB 2;	Length 2277;	0;
Best Local Similarity	27.4%;	Pred. No. 1.5;			
Matches 80;	Conservative 50;	Mismatches 162;	Indels 0;	Gaps 0;	
Qy	86	ggagggccatgaggggctgaacgtcacagcaacgagcgagggcctgccacgcccttt	145		
Db	503	GNAARGCNYTNAARGGNATGAAYATGAARGARATHAARGNMGNACNGTNGCNGTNGAYT	562		
Qy	146	atttgcctctacgaaaaagttaaaaaagacattgagtgatgtaatccacctgggggca	205		
Db	563	GGCGCCTTNGCNAARGAAYARTYAAARGAYACNCARWSNGTWNWSGNCNATHGGNGARGARA	622		
Qy	206	atagcccatattgccaatgggtgcggccgggtgtgtggcaacattacttcatgtagcacc	265		
Db	623	ARWSNCAYARWSNAARCAVCARGARWSNGTNAARAARAARGGNMNGARGARGARGAYA	682		
Qy	266	tgaacctgggaaggtaatgattctccacctatacctctgtggcagctgcaacctgta	325		
Db	683	TGGARGARGARGARAAAYGAYGAYGAYGAYGAYGAYGARGARGAYGCGGNTNTTYG	742		
Qy	326	ttctttcacagtttcgagaagaagcacatcagttttgggtgggaatactgt	377		
Db	743	AYGAYGARGAYGARGARGARGARAAYATHGARWWSNAARGTWNACNAARCCNT	794		

```

RESULT      8
US-08-630-915A-33
: Sequence 33, Application US/08630915A
: Patent No. 6309820
: GENERAL INFORMATION:
: APPLICANT: SPARKS, Andrew B.
: APPLICANT: HOFFMAN, No. 6309820h
: APPLICANT: KAY, Brian K.
: APPLICANT: FOWLKES, Dana M.
: APPLICANT: MCCONNELL, Stephen J.
: TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
: TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF
:

```



```
;
; TITLE OF INVENTION: USING SAME
; NUMBER OF SEQUENCES: 227
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/630,915A
; FILING DATE: 03-APR-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: MISTOCK, S. LESLIE
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-174
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8664/9741
; TELEX: 66141 PENNIE
;
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4091 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
;
; US-08-630-915A-33
;
; Query Match 7.5%; Score 29.6; DB 4; Length 4091;
; Best Local Similarity 51.5%; Pred. No. 3.2;
; Matches 68; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
;
; Qy 10 ctacagctgacccagctgcccgtatcgcaatggtgtgagggccctctggaggattata 69
; Db 995 CGAGACTGGAGCAGGCCATCCGGGGCCGATGCCAGGAGGACCTCAGGTGGTTCGC 1054
;
; Qy 70 agaagcggggccctatgaggcccatgaggggctgaaactcacagcaacagcgccaggg 129
; Db 1055 AGCACAGTGGCCCGGGATGCCATGAACCTGCGCCAGTTCGAGAGTGGAAACCAGAC 1114
;
; Qy 130 cctgcccacgcc 141
; Db 1115 CTCCCGCACACC 1126
;
; RESULT 9
; US-08-475-035-3
; Sequence 3, Application US/08475035
; Patent No. 5985553
; GENERAL INFORMATION:
; APPLICANT: KING, C. R.
; APPLICANT: KRAUS, MATTHIAS H.
; APPLICANT: AARONSON, STUART A.
; TITLE OF INVENTION: HUMAN GENE RELATED TO BUT DISTINCT FROM
; TITLE OF INVENTION: EGF RECEPTOR GENE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG, P.C.
; STREET: Suite 1200, 127 peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
;
; TITLE OF INVENTION: USING SAME
; NUMBER OF SEQUENCES: 227
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,035
; FILING DATE: 7 Jun 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G.
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.656
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404/688-0770
; TELEFAX: 404/688-9880
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5532 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 187..3816
;
; US-08-475-035-3
;
; Query Match 7.4%; Score 29.2; DB 2; Length 5532;
; Best Local Similarity 53.5%; Pred. No. 5.2;
; Matches 61; Conservative 0; Mismatches 53; Indels 0; Gaps 0;
;
; Qy 258 tgacgccatgaacctgcccgaaggaatttcctcaacctatcacctctgtggcgagctg 317
; Db 4605 TGCTGTGTAATAATCAGCAAGAGAGATGACACATCAATAATACTCGGATTCAGCCC 4664
;
; Qy 318 cacctgtattttttacagtttgcagaagaaagcacatcacagtttgggtgggaaa 371
; Db 4665 ACATTGGATTTCAGCATTTGGACCAATAGCCACAGCTGAGAATGTGGAATA 4718
;
; RESULT 10
; US-08-202-389-1
; Sequence 1, Application US/08202389
; Patent No. 5536636
; GENERAL INFORMATION:
; APPLICANT: FREEMAN JR., ROBERT M.
; APPLICANT: PLUTZKY, JORGE
; APPLICANT: NEEL, BENJAMIN G.
; APPLICANT: ROSENBERG, ROBERT D.
; TITLE OF INVENTION: IDENTIFICATION OF NOVEL TYROSINE
; TITLE OF INVENTION: PHOSPHATASES HAVING SH2 DOMAINS
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/202,389
; FILING DATE: 28-FEB-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/983,926
; FILING DATE: 01-DEC-1992
; APPLICATION NUMBER: US 07/829,141
```

; FILING DATE: 31-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/721,112
; FILING DATE: 26-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: BH92-05WA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1747 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..1540
US-08-202-389-1

Query Match 7.3%; Score 29; DB 1; Length 1747;
Best Local Similarity 57.0%; Pred. No. 3.3;
Matches 53; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
Qy 96 gagggggtgaacgtcacagcaaacagggcgagggcgctgcacgccctttatttgctg 155
Db 105 GAGGCGAGCGACAGTCACTGTCAGGCCAACAGGGCGAGCCCTGGACATTTCTGTGCGTG 164
Qy 156 ctacgaaagttaaaaaagacattgagtgt 188
Db 165 AGAGTCTCAGCAACCTGCTGATTTGTGCTCT 197

RESULT 11
US-09-457-708-1/c
; Sequence 1, Application US/09457708
; Patent No. 6326483
; GENERAL INFORMATION:
; APPLICANT: Kwiatkowski, David J.
; APPLICANT: Sampson, Julian R.
; APPLICANT: Povey, Sue
; APPLICANT: van Slegtenhorst, Marjon
; APPLICANT: Halley, Dicky
; TITLE OF INVENTION: Compositions and Methods Based U
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Vinson & Elkins
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.
; ZIP: 20004-1008
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/457,708
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sanzo, Michael A.
; REGISTRATION NUMBER: 36,912
; REFERENCE/DOCKET NUMBER: BRI331/42002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 639-6585
; TELEFAX: (202) 639-6604

; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8600 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-09-457-708-1

Query Match 7.3%; Score 29; DB 4; Length 8600;
Best Local Similarity 58.8%; Pred. No. 7.7;
Matches 50; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
Qy 100 gggctgaacgtcacagcaaacagggcgagggcgctgcacgccctttatttgctgctac 159
Db 7750 GGGACAAGGGTCACAGCAGCAGCCCTAGGGGCGCAGTCTCGACCTAAACTTGTCTTTCAC 7691
Qy 160 gaaaagttaaaaaagacattgagtgt 184
Db 7690 CTACAGACAAAAGCTTAATCAAGTG 7666

RESULT 12
US-09-073-569-4/c
; Sequence 4, Application US/09073569
; Patent No. 6084088
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Grossmann, Angelika
; TITLE OF INVENTION: NOVEL TUMOR ANTIGENS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,569
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sawislak, Deborah A
; REGISTRATION NUMBER: 37,438
; REFERENCE/DOCKET NUMBER: 97-14
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6672
; TELEFAX: 206-442-6678
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1311 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-09-073-569-4

Query Match 7.2%; Score 28.6; DB 3; Length 1311;
Best Local Similarity 30.7%; Pred. No. 3.9;
Matches 42; Conservative 25; Mismatches 70; Indels 0; Gaps 0;

Qy	170	aaagacattgagtgtataatccaccctgggggcaatatgccatatcgccaatggtggc	229
		: : : : : : : : : : :	
Db	626	ARRAATCYTTCNKTTRCAATTTCNGTCATNCCNACNGGCCDATYTCYTGGTGNCGR	567
		: : : : : : : : : : :	
Qy	230	ccgggtgtggaacaattacttcattgatgcagccatgaacctgcggaagtgtaattgatt	289
		: : : : : : : : : : :	
Db	566	TTNARNARTTRCANCNGSYTGGGCATRCACCCTGNACNCKNARRTTNSWRAADATN	507
		: : : : : : : : : : :	
Qy	290	cctcaaacctcatcactct 306	
		: : : : : : : :	
Db	506	CCNCCNCCNCKNARNCK 490	

```

RESULT 13
US-08-026-138E-17/c
; Sequence 17, Application US/08026138E
; Patent No. 5502166
; GENERAL INFORMATION:
; APPLICANT: Masayoshi MISHINA
; TITLE OF INVENTION: NOVEL PROTEINS AND GENES CODING THE SAME

```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match 7.2%; Score 28.4; DB 1; Length 4368;
Best Local Similarity 70.4%; Pred. NO. 8.7;
Matches 38; Conservative 0; Mismatches 16; Indels 0

Qy 64 attataaacgagggcctatgaggcccatgagggggctgaacgcacagca 117
||||| ||| || ||||| ||||| || ||||| ||| ||| |||||
Db 323 ATTATCATAGATGAGCCCCCATGGATGCCAGGATGGGGGTGAGAGTCTGAGCA 270

RESULT 14
US-08-026-138E-6/c
; Sequence 6, Application US/08026138E

Patent No. 5502166
GENERAL INFORMATION:
APPLICANT: Masayoshi MISHINA
TITLE OF INVENTION: NOVEL PROTEINS AND GENES CODING
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nishiohata Residence 1-107
STREET: 5214, Nishiohata-machi

Query Match 7.2%; Score 28.4; DB 1; Length 4446;
Best Local Similarity 70.4%; Pred. No. 8.8;
Matches 38; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 64 attatagaacggagggcctatggaggcccatgaggggctgaacgtcacagca 117
 ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 401 ATTATCATAGTAGACCCCCCATGGATGCCAGGATGGGGGTGAGGTCTGAGCA 348

RESULT 15
US-09-268-163-9/c

```
; Sequence 9, Application US/09260163B
; Patent No. 6353091
; GENERAL INFORMATION:
; APPLICANT: Lipscombe, Diane
; APPLICANT: Schorge, Stephanie
; TITLE OF INVENTION: HUMAN N-TYPE CALCIUM CHANNEL ISOFORM AND USES THEREOF
; FILE REFERENCE: B1055/7000
; CURRENT APPLICATION NUMBER: US/09/268,163B
; CURRENT FILING DATE: 1999-03-12
; EARLIER APPLICATION NUMBER: US 60/077,901
; EARLIER FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 7011
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..7008
US-09-268-163-9
```

```
Query Match      7.1%; Score 28.2; DB 4; Length 7011;
Best Local Similarity 54.3%; Pred. No. 13;
Matches 57; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 167 taaaaagacattgagtgaatccaccctgggggcaatagccatatigccaatggtg 226
      ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6357 TGAAGAGAGAGAGGGGCTCCCTCCGCTCTCTGGGACCGGCTCGCTCTGCTTACGCTC 6298

QY 227 cggccgggtgtgtggcaacattacttcattgatgcagccatgaacc 271
      ||||| ||| ||| ||| ||||| ||||| ||||| ||||| |||||
Db 6297 CCGCCGCGAGCCTGTGGACCCTTCTCCATGGGGGCGAGCCAGATCC 6253
```

Search completed: September 28, 2002, 07:56:08
Job time: 11253 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 28, 2002, 06:22:49 ; Search time 4881.77 Seconds
(without alignments)
1094.848 Million cell updates/sec

Title: us-09-870-113-11

Perfect score: 396

Sequence: 1 atcagagttacagctga.....ttacctgtttaccactag 396

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	280	70.7	747	10	BI827322 603077989
2	277.6	70.1	888	9	AL530804 603077989
3	265.8	67.1	480	10	BI045863 MR3-FN020
4	256	64.6	573	10	BI541703 455541 MA
5	246.4	62.2	598	10	BE913718 601669072
6	246.4	62.2	645	10	BG964218 602828996
7	246.4	62.2	1131	10	BG295496 602392826
8	244.8	61.8	560	10	BI794616 1c92e02.y
9	244.6	61.8	574	10	BI185580 UNT-P-FN-
10	234	59.1	501	9	AW211366 uo80b07.y
11	224.4	56.7	701	10	BI831943 603076873
12	223.4	56.4	336	9	AW326482 19106 MAR
13	222.8	56.3	817	10	BI103329 602888838
14	195.4	49.3	546	10	BM488747 pgm2n.pk0
15	191.6	48.4	453	10	BF916224 CM2-UT011
16	191.4	48.3	855	10	BI412175 602965377
17	188.4	47.6	967	10	BI739710 603361983

```

18 185.6 46.9 397 9 BE012485
19 182.2 46.0 697 10 BI854643
20 168 42.4 722 10 BI876051
21 166.2 42.0 622 10 BJ039504
22 166 42.0 681 9 AV704087
23 156.8 39.6 421 10 BI448355
24 149.8 37.8 637 9 AL654963
25 143.8 36.3 417 9 AA104365
26 142.6 36.0 546 9 BE014800
27 140.6 35.6 561 9 AI942584
28 139.4 35.2 678 9 BB654057
29 139.2 35.2 1295 11 AK015770
30 137.8 34.8 916 10 BI757180
31 129.6 32.7 567 10 BJ035232
32 129.6 32.7 612 10 BJ012200
33 128.2 32.4 512 9 AW210205
34 128 32.3 684 9 AI133696
35 128 32.3 750 10 BI831750
36 126.4 31.9 488 9 AI310713
37 126 31.8 309 9 AI630395
38 125.4 31.7 729 10 BI460337
39 116 29.3 803 10 BG777404
40 115.2 29.1 1209 11 AK006155
41 113.8 28.7 331 10 BI449976
42 109.8 27.7 666 9 BB644566
43 107.6 27.2 593 10 BG895490
44 104.8 26.5 483 9 AA061624
45 103.4 26.1 350 12 AZ693917

```

ALIGNMENTS

RESULT 1

BI827322

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

```

BI827322      747 bp  mRNA  linear  EST 04-OCT-2001
603077989F1  NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5169694 5',
mRNA sequence.
BI827322
BI827322.1  GI:15938872
EST
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/
1 (bases 1 to 747)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Prepared by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11422 row: a column: 23
High quality sequence stop: 743.
Location/Qualifiers
1. 747
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5169694"
/clone_lib="NIH_MGC_119"
/lab_host="DH10B"

```

```

Site_1: Notif;
Site_2: EcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb, insert size range

```

FEATURES
source

0.9-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 013. Note: this is a NIH_MGC Library.*

BASE COUNT 188 a 206 c 201 g 151 t 1 others
ORIGIN

Query Match 70.7%; Score 280; DB 10; Length 747;
Best Local Similarity 100.0%; Pred. No. 6.2e-78;
Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atgcagagctctacagcctgaccagctgccgctatcgcaatgtgttgagccctctgg 60
|||||
Db 84 ATGCAGAGCTTACAGCCTGACCCAGCTGCCGCTATCGCAATGTGTGGAGGCCCTCTGG 143
Qy 61 aggattataagaacaggagggccctatggggcccatgaggggctgaaacgtcacagcaaca 120
|||||
Db 144 AGGATTATAAGAACAGGAGGGCCTATGGAGGCCCATGAGGGGCTGAACGTCACAGCAACA 203
Qy 121 ggcgcagggcctgccagcgcctttatttgcctgctacgaaaaagttaaaaaagacattg 180
|||||
Db 204 GCGCAGGGCCTGCCAGCGCCCTTTATTTTGCCTCTACGAAAAGTTAAAAAGACATTG 263
Qy 181 agtgatgtaatccaccctgggggcaatagccatattgccatattgctgcggcggtgtgtg 240
|||||
Db 264 ACTGATGTAATCCACCCCTGGGGCAATAGCCATATTGCCAATGTGCGCGGGGTGTGTG 323
Qy 241 gcaacattacttcattgatgcagccatgaacccctgcggaag 280
|||||
Db 324 GCAACATTACTTTCATGTCAGCCCATGAACCCCTGCGGAAG 363

RESULT 2
AL530804
LOCUS AL530804 LTI_NFI001_NBC4 888 bp mRNA linear EST 13-FEB-2001
DEFINITION prime, mRNA sequence.
ACCESSION AL530804
VERSION AL530804.1 GI:12794297
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Li.W.B., Gruber.C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1. .888
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSDD008VH17"
/clone_lib="LTI_NFI001_NBC4"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with NotI and cloned into the NotI and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 168 a 244 c 322 g 139 t 15 others

ORIGIN

Query Match 70.1%; Score 277.6; DB 9; Length 888;
Best Local Similarity 99.3%; Pred. No. 3.8e-77;
Matches 278; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 atgcagagctctacagcctgaccagctgccgctatcgcaatgtgttgagccctctgg 60
|||||
Db 417 ATGCAGAGCTTACAGCCTGACCCAGCTGCCGCTATCGCAATGTGTGGAGGCCCTCTGG 476
Qy 61 aggattataagaacaggagggccctatggggcccatgaggggctgaaacgtcacagcaaca 120
|||||
Db 477 AGGATTATAAGAACGGAGGGCCTATGGAGGCCCATGAGGGGCTGAACGTCACAGCAACA 536
Qy 121 ggcgcagggcctgccagcgcctttatttgcctgctacgaaaaagttaaaaaagacattg 180
|||||
Db 537 GCGCAGGGCCTGCCAGCGCCCTTTATTTTGCCTCTGCBGAAAAAGTTAAAAAGACATTG 596
Qy 181 agtgatgtaatccaccctgggggcaatagccatattgccatattgctgcggcggtgtgtg 240
|||||
Db 597 AGTATGTTATCCACCCCTGGGGCAATAGCCATATTGCCAATGTGCGCGGGGTGTGTG 656
Qy 241 gcaacattacttcattgatgcagccatgaacccctgcggaag 280
|||||
Db 657 GCAACATTACTTTCATGTCAGCCCATGAACCCCTGCGGAAG 696

RESULT 3
BI045863
LOCUS BI045863 BR3-FN0209-300101-004-h10 FN0209 Homo sapiens cDNA, mRNA sequence.
DEFINITION
ACCESSION BI045863
VERSION BI045863.1 GI:14452485
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,F., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR3&t2=MR3-FN0209-300101-004-h10&t3=2001-01-30&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 480.
FEATURES
source
1. .480
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="FN0209"
/dev_stage="Adult"
/note="Organ: prostate_normal; Vector: puc18; Site_1: SmaI ; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent

application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT
ORIGIN

116 a 126 c 146 g 92 t

Query Match 67.1%; Score 265.8; DB 10; Length 480;
Best Local Similarity 98.9%; Pred. No. 1.8e-73;
Matches 278; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 atgcagagctcacagcctgaccagctgcccgcctatcgcaatgtgttgagggccctctgg 60
Db 74 ATGCAGAGCTTACAGCCTACACCGCTGCCGCTATCGCAATGTGTGAGGCCCTCTGG 133
QY 61 aggattataagaacgagggcctatggagcccatgagggcgtgaacatcacacaaca 120
Db 134 AGGATTATTAAGAACGGAGGGCCTATGGAGGCCATGAGGGGCTGAACGTCACAGCAACA 193
QY 121 ggcgagggcctgcccacgcccctttatttgcctgctacgaaagttaaaaaagacattg 180
Db 194 GCGCAGGGCCCTGCCACCGCTTATTATTGCTGCTACGAAAGTTAAAAAGACATTG 253
QY 181 agtgatgtaataccacct-gggggcaatagccatattgccaatgtgctgcccgggtgtgt 239
Db 254 AGTGATGTAATCCACCTGGGGGGAATAGCCATATTGCCAATGTGTGCGCGGGTGTGT 313
QY 240 gcaacattactctatgatgcagccatgaacctgcggaag 280
Db 314 GGCACATTACTTCATGATGATGCAGCCATGAACCTGCGGAAG 354

RESULT 4

LOCUS B1541703 573 bp mRNA linear EST 30-AUG-2001
DEFINITION 455541 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION B1541703
VERSION B1541703.1 GI:15382815
KEYWORDS EST.
SOURCE COW.

ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 573)
Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett
G.L., Heaton,M.P., Laegreid,W., Rohrer,G.A., Chitko-McKown,C.G.,
Perteaux,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and
Keele,J.W.

TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle

JOURNAL Genome Res. 11 (4), 626-630 (2001)
MEDLINE 21180013
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@mail.marc.usda.gov

Single pass sequencing. Bases called and alt_trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.

PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCTCCAGTCAGCAG
Plate: 121 row: O column: 24
Seq primer: ATTTAGGTGACACTATAG.

FEATURES

Source
1..573
Location/Qualifiers
/organism="Bos taurus"
/db_xref="taxon:9913"

/clone_lib="MARC 1BOV"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled tissue from lymph node, ovary,
fat, hypothalamus, and pituitary."

BASE COUNT 121 a 166 c 176 g 110 t
ORIGIN

Query Match 64.6%; Score 256; DB 10; Length 573;
Best Local Similarity 94.6%; Pred. No. 2.4e-70;
Matches 265; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 atgcagagctcacagcctgaccagctgcccgcctatcgcaatgtgttgagggccctctgg 60
Db 174 ATGCAGAGCTTACAGCGGATCCAGCCGCCCTATCGCAATGTGTGAGGCCCTCTGG 233
QY 61 aggattataagaacgagggcctatggagggcccatgaggggctgaacctcacacaaca 120
Db 234 AGGATTATTAAGAACGGAGGGCCTGTGGAGGCCATGCGAGGCTGAACGTCACAGCAACA 293
QY 121 ggcgagggcctgcccacgcccctttatttgcctgctacgaaagttaaaaaagacattg 180
Db 294 GCGCAGGGCCCTGCCACCGCCCTCTATTGCTGCTATGAAAAGTTAAAAAGACATTG 353
QY 181 agtgatgtaataccacctgggggcaatagccatattgccaatgtgctgcccgggtgtgtg 240
Db 354 AGTGATGTAATCCACCTGGGGGCAATAGCCATATTGCCAATGTGTGCGCGGTGTGTG 413
QY 241 gcaacattactctatgatgcagccatgaacctgcggaag 280
Db 414 GCAACATTACTTCATGATGCTGCCATGAATCCAGTGGGAAG 453

RESULT 5

LOCUS BE913718 598 bp mRNA linear EST 29-SEP-2000
DEFINITION 601669072F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:3969256 5',
mRNA sequence.

ACCESSION BE913718
VERSION BE913718
KEYWORDS EST.
SOURCE house mouse
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 598)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Plate: LLAM9146 row: o column: 17
High quality sequence stop: 593.

FEATURES

Source
1..598
Location/Qualifiers

/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:3969256"
/clone_lib="NCI_CGAP_Mam1"
/tissue_type="tumor, biopsy sample"
/dev_stage="10 months, virgin"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;

[illegible]

RESULT	8	
LOCUS	BI794616	
DEFINITION	ic92e02.y1 Melton Normalized Mixed Mouse Pancreas 1 N1-MMS1 Mus musculus cDNA 5' similar to TR:Q23125 Q23125 W02B12.9 PROTEIN. [1]	
	; mRNA sequence.	
ACCSSION	BI794616	560 bp mRNA linear EST 01-OCT-2001
VERSION	BI794616.1	GI:15822341
KEYWORDS	house mouse.	
SOURCE	Mus musculus	
ORGANISM	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
REFERENCE	1 (bases 1 to 560)	
AUTHORS	Melton, D., Brown, J., Keny, G., Permutt, A., Lee, C., Kaestner, K., Lemishka, I., Secorce, M., Brestelli, J., Gradwohl, G., Clifton, S., Hillier, L., Marra, M., Pape, D., Wyllie, T., Martin, J., Blisstein, A., Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagareishvili, R., Williams, T., Jackson, Y. and Bowers, Y.	
TITLE	Endocrine Pancreas Consortium	
JOURNAL	Unpublished (2000)	
COMMENT	Other ESTs: ic92e02.x1	

Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
 Endocrine Pancreas Consortium
 Harvard University, Howard Hughes Medical Institute
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
 MA 02138
 Tel: 617-495-1812
 Fax: 617-495-8557
 Email: dmelton@biohp.harvard.edu
 Library was constructed by Dr. Douglas Melton DNA sequencing by:
 Washington University Genome Sequencing Center For information on
 obtaining a clone please contact: Juliana Brown
 (brown@fas.harvard.edu)

```
Seq primer: -40RP from Gibco
High quality sequence stop: 430.
Location/Qualifiers
1..560
/organism="Mus musculus"
/strain="ICR"
/db_xref="taxon:10090"
/clone_lib="Melton Normalized Mixed Mouse Pancreas 1
NI-MMS1"
/sex="Both for embryonic & newborn, male for adult and
adult islet"
/dev_stage="Embryonic day 10.5, E12.5, E16.5, newborn,
adult, mixed"
/lab_host="DH10B"
```

/note="Vector: pSPOR1; Site_1: Not I; Site_2: Sal I; Five libraries representing E10.5/12.5 pancreatic bud, E16.5 pancreas, newborn pancreas, adult pancreas, and adult islets of Langerhans were separately constructed using SuperScript Plasmid library kit (Life Technologies). cDNA was made by oligo-dT priming and size-selected by column fractionation. Libraries were amplified once on solid support and plasmid DNA from each library was prepared and mixed in equal amounts. The mixed library DNA was normalized by method #4 from Bonaldo, Lennon, and Soares 1996 Genome Research 6:791-806; 0.5 microgram single-stranded mixed library plasmid DNA was mixed with 5 micrograms PCR product representing mixed library inserts and hybridized to an EcoT of 6. Single-stranded (hybridized) plasmids were isolated by hydroxyapatite chromatography and used to make this library."

BASE COUNT 134 a 158 c 163 g 105 t
ORIGIN

Query Match 61.8%; Score 244.8; DB 10; Length 560;
Best Local Similarity 92.1%; Pred. No. 8.8e-67;
Matches 258; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy	1	atgcagagctacagcctgaccagctgcccgctatcgcgaatgtgtggaggccctctgg	60
Db	84	ATGCAGAGCCTACAGCCTGACCCAGCGCCCGCTATCGGAACGTGTGGAGGCTCTCTGG	143
Qy	61	aggattataagaacgagagccctatgagagcccatgaggggcctgaactcacagcaaca	120
Db	144	AGATCATGAGGACAGAGGCGCTGTGGAGGCCCATCGGGGGCTGACCGTACAGTAACA	203
Qy	121	ggcgagggcctgcccacgcccctttatttgcctgctacgaaagttaaaaagacatcg	180
Db	204	GGCGGGGGCTGCCACGCCCTCTATTTTGCCTGTCTACGAAAGTTAAAAAGACATTG	263
Qy	181	agtgatataacacctgggggcaatagccatatgtccaatggtcgccggggtgtgtg	240
Db	264	AGTCACTTAATCCACCCAGGGGGCAATAGCCATATTGCCAATGGTGCAGCCGATGTGTG	323
Qy	241	gcaacattactctatgatgcagcctagaacccctgcggaag	280
Db	324	CGGACATTACTTCATGATCAGGCATGAATCCACGCGAAG	363

RESULT	9
LOCUS	BI185580
DEFINITION	BI185580 UNL-P-FN-cy-f-02-0-UNL.s1 UNL-P-FN Sus scrofa cDNA clone
ACCESSION	BI185580
VERSION	BI185580.1
KEYWORDS	EST.
SOURCE	plq.
	574 bp mRNA linear EST 10-JUL-2001

Org. Sus. scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 1 (bases 1 to 574)
 Caetano,A.R., Johnson,R.K. and Pomp,D.
 Generation and sequence characterization of a normalized cDNA
 library from swine ovarian follicles
 Unpublished (2001)
 Contact: Pomp, D
 Department of Animal Science
 University of Nebraska, Lincoln
 Lincoln, NE 68583-0908, USA
 Tel: 402 472 6416
 Tel: 402 472 6416
 Fax: 402 472 6362
 Email: dpomp@unl.edu
 Oligo-dT track not found, Not I site shown in beginning of sequence
 is likely internal to the message.
 Seq primer: M13 -29
 POLYA=NO.

cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

Plate: L1AM11419 row: a column: 20
 High quality sequence start: 29
 High quality sequence stop: 699.

FEATURES

source
 1. 701
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5168539"
 /clone_lib="NIH_MGC_119"
 /tissue_type="medulla"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: EcoRV (destroyed); RNA source normal medulla from
 anonymous male age 27. Library is oligo-dT primed and
 directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.3 kb, insert size range
 0.9-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 013. Note:
 this is a NIH_MGC Library."
 BASE COUNT 176 a 167 c 208 g 150 t
 ORIGIN

Query Match 56.7%; Score 224.4; DB 10; Length 701;
 Best Local Similarity 99.6%; Pred. No. 2.9e-60;
 Matches 225; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 atgcagagctcacagcctgacccagctccgctatcgcaatgtgttgaggccctctgg 60
 Db 302 ATGCAGAGCTTACAGCTTACCCAGCTCCGCTATCCATGTTGTGGAGGCCCTCTGG 361
 Qy 61 aggtattataagaagcggagggcctatggaggcccatgagggcctgaaagctcacagcaaca 120
 Db 362 AGGATTATTAAGAAGAGGAGGGCTATGAGAGGCCCATGAGGGGCTGAAGCTCACAGCAACA 421
 Qy 121 ggcagggcctgcccagccctttatttgcctgctacgaaagttaaaaagacattg 180
 Db 422 GCGCAGGGCCTGCCAGCCCTTATTTCCTGCTACGAAAAGTTAAAAAGACATTG 481
 Qy 181 agtgatgaatccaccctgggggcaatagccatatgtgccaatggtg 226
 Db 482 AGTGATGTAATCCACCCCTGGGGGCANTAGCCATATTGCCAATGCTG 527

RESULT 12
 AW326482
 LOCUS 336 bp mRNA linear EST 25-APR-2001
 DEFINITION 19106 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.
 ACCESSION AW326482
 VERSION AW326482.1 GI:6762403
 EST.
 KEYWORDS EST.
 SOURCE cow.

ORGANISM Bos taurus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.
 REFERENCE 1 (bases 1 to 336)
 AUTHORS Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
 Casas,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett
 ,G.L., Eaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-Mckown,C.G.,
 Pertea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and
 Keele,J.W.

TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA
 libraries and construction of a gene index for cattle
 JOURNAL Genome Res. 11 (4), 626-630 (2001)
 MEDLINE 21180013

COMMENT

Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smith@email.marc.usda.gov
 Single pass sequencing. Bases called and trimmed with phred
 v0.980904.e. Vector identified by cross_match with the -minscore 20
 and -minmatch 12 options.

PCR Primers
 FORWARD: AGGAACAGCTATGACCAT
 BACKWARD: GTTTCCTCAGTCACGACG
 Plate: 8 row: F column: 16
 Seq primer: ATTTAGGTGACACTATAG.

FEATURES

source
 1. 336
 /organism="Bos taurus"
 /db_xref="taxon:9913"
 /clone_lib="MARC 1BOV"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
 Library made from pooled tissue from lymph node, ovary,
 fat, hypothalamus, and pituitary."
 BASE COUNT 72 a 96 c 107 g 61 t
 ORIGIN

Query Match 56.4%; Score 223.4; DB 9; Length 336;
 Best Local Similarity 95.4%; Pred. No. 4.7e-60;
 Matches 230; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
 Qy 1 atgcagagctcacagcctgacccagctccgctatcgcaatgtgttgaggccctctgg 60
 Db 96 ATGCAGAGCTTACAGCTTACCCAGCTCCGCTATCCATGTTGTGGAGGCCCTCTGG 155
 Qy 61 aggtattataagaagcggagggcctatggaggcccatgagggcctgaaagctcacagcaaca 120
 Db 156 AGGATTATTAAGAAGAGGAGGGCTGTGGAGGCCCATGCGAGGCTGAGCGCTCACAGCAACA 215
 Qy 121 ggcagggcctgcccagccctttatttgcctgctacgaaagttaaaaagacattg 180
 Db 216 GCGCAGGGCCTGCCAGCCCTTATTTCCTGCTATGAAAAGTTAAAAAGACATTG 275
 Qy 181 agtgatgaatccaccctgggggcaatagccatatgtgccaatggtgcccgggtgtgtg 240
 Db 276 AGTGATGTAATCCACCCCTGGGGGCANTAGCCATATTGCCAATGTCGGCTGGTGTGTG 335
 Qy 241 g 241
 Db 336 G 336

RESULT 13
 B1103329
 LOCUS 817 bp mRNA linear EST 26-JUN-2001
 DEFINITION 60288838F1 NCI_CGAP_K1d14 Mus musculus cDNA clone IMAGE:5043992
 5', mRNA sequence.
 ACCESSION B1103329
 VERSION B1103329.1 GI:14554222
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 817)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM1120 row: 1 column: 09
 High quality sequence stop: 777.

FEATURES
 source
 Location/Qualifiers
 1..817

BASE COUNT 147 a 222 c 316 g 132 t
 ORIGIN
 Query Match 56.3%; Score 222.8; DB 10; Length 817;
 Best Local Similarity 91.5%; Pred. No. 9.9e-60;
 Matches 258; Conservative 0; Mismatches 22; Indels 2; Gaps 2;
 QY 1 atcgaggtctacagctgacccagctgcccgtatgcacgtatggtgagccctctgg 60
 Db 422 ATGCAGAGCTTACAGCTGACCCAGCCGCCGCTATCGGAACGTGTGGAGGCTCTCG 481
 QY 61 aggattataagaacggaggcctatgagggccatgagggcctgaaacgtcacagcaaca 120
 Db 482 AGAATCATGAGACAGAGGGGCTCTGGAGGCCCATGCGGGGCTGAACGTACAGCAACA 541
 QY 121 ggcgagggcctgccacgccccttatttgcctgctacgaaagttataaagacattg 180
 Db 542 GGCAGGGGCTGCGCCAGCCCTCTATTTGCTCTACGAAAAGTTAAAAAGACATTG 601
 QY 181 agtataatccaccctggggcgaatagccattatccattg-tgcggccggtgtgt 239
 Db 602 AGTACGTAATCCACCCAGGGGGCAATAGCCATATTCCTATGTTGACGCGGATGTGT 661
 QY 240 ggcacattac-ttcattgatgcagccatgaacccctgcggaag 280
 Db 662 GCGACATTAATTCATGATGCAGGCATGAATCCAGCGGAAG 703

RESULT 14
 BM488747
 LOCUS
 DEFINITION
 pgm2n.pk008.12 Normalized Chicken Breast Muscle, Leg Muscle, and
 Epiphyseal Growth Plate cDNA library (pgm2n) Gallus gallus cDNA
 clone pgm2n.pk008.12 5' similar to ref:XP_050766.1 (XM_050766)
 putative mitochondrial solute carrier [Homo sapiens]
 gb|AAK49519.1|AF327402.1 (AF327402) putative mitochondrial solute
 carrier splice variant [Homo sapiens], mRNA sequence.
 ACCESSION
 BM488747
 VERSION
 BM488747.1 GI:18609678
 KEYWORDS
 EST.
 SOURCE
 chicken.
 ORGANISM
 Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 REFERENCE
 1 (bases 1 to 546)
 Cogburn,L.A. and Monsonego-Ornan,E.
 ESTs from Normalized Chicken Breast
 Epiphyseal Growth Plate cDNA library, USDA/IRAFS Animal Genome
 Project.
 JOURNAL
 Unpublished (2002)
 COMMENT
 Contact: Larry A. Cogburn
 University of Delaware

Townsend Hall, Newark, DE 19717, USA
 Tel: 302-831-1335
 Fax: 302-831-2822
 Email: cogburn@udel.edu, www.chickest.udel.edu.
 Location/Qualifiers
 1..546

FEATURES
 source
 Location/Qualifiers
 1..546

BASE COUNT 90 a 183 c 197 g 70 t 6 others
 ORIGIN
 Query Match 49.3%; Score 195.4; DB 10; Length 546;
 Best Local Similarity 80.7%; Pred. No. 4.5e-51;
 Matches 226; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
 QY 1 atcgaggtctacagctgacccagctgcccgtatgcacgtatggtgagccctctgg 60
 Db 217 ATGCAGAGCTTGGGCCCGAGCCGCCGCTACCGAACGTGTGGAGGCTCTGG 276
 QY 61 aggattataagaacggaggcctatgagggccatgagggcctgaaacgtcacagcaaca 120
 Db 277 CGCATCTGCTACCGAGGGCGTGTGGAGGCCCATGCGGGGCTGAACATCACGCCACC 336
 QY 121 ggcgagggcctgccacgccccttatttgcctgctacgaaagttataaagacattg 180
 Db 337 GGCAGGGGCTGCGCCAGCCCTCTACTTGCCTGTACGAAAAGTTAAAAAGACGCTG 396
 QY 181 agtataatccaccctggggcgaatagccattatccattg-tgcggccggtgtgt 240
 Db 397 AGCAGCTCATCCACGGGGGGCAATAGCCATGTGCGNNCGTGCGAGCGGGGTGTGA 456
 QY 241 gcaacattacttcattgatgcagccatgaacccctgcggaag 280
 Db 457 GCAACATTGCTCCACGACGCGCATGCAACCCCTGCAGAAG 496

RESULT 15
 BF916224
 LOCUS
 DEFINITION
 CM2-UT0112-091200-598-c06 UT0112 Homo sapiens cDNA, mRNA sequence.
 ACCESSION
 BF916224
 VERSION
 BF916224.1 GI:12307682
 KEYWORDS
 EST.
 SOURCE
 human.
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 453)
 Dias Neto,E., Garcia Correa,R., Verjowski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
 ,M.J., Soares,F., Brentani,R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags

THIS PAGE BLANK (USPTO)